

Gencore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using fframe_plus.p2n model

Run on: February 26, 2003, 19:53:56 ; Search time 3193 Seconds

(without alignments)
5897.123 Million cell updates/sec

Title: US-09-677-653a-50
Perfect score: 3374
Sequence: 1 MDDAGVASRPNHRRCTRVN.....GRIARVRRARRRRARRAN 647

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=framed.p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09677653/runat_20022003_164315_1251/app_query.fasta.1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09677653.ecgn_1_1_2034.etrnat_20022003_164315_1251 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_dat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_iny:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3374	100.0	2478	6 AR125096	AR125096 Sequence
2	3374	100.0	2478	6 AR125097	AR125097 Sequence
3	3374	100.0	2478	6 AR159968	AR159968 Sequence
4	3374	100.0	2478	6 AX028950	AX028950 Sequence
5	3374	100.0	2478	6 AX028971	AX028971 Sequence
6	3374	100.0	2478	14 HVS00AT	L37299 Helicoverpa
7	3361	99.6	2479	6 AR125098	AR125098 Sequence
8	2207	65.4	2482	14 S43937	S43937 capsid prot
9	436	12.9	2482	14 AF062037	AF062037 Thosaa as
10	407.5	12.1	5638	14 AF461742	AF461742 Euproster
11	386.5	11.5	6625	14 AF102884	AF102884 Nudaurell
12	294	8.7	6534	6 AR159969	AR159969 Sequence
13	294	8.7	6536	6 AX028953	AX028953 Sequence
14	294	8.7	6536	6 AX028972	AX028972 Sequence
15	155.5	4.6	11659	1 AE005131	AE005131 Halobacte
16	153	4.5	151826	2 AC125378	AC125378 Mus muscu
17	148.5	4.4	110000	2 LMFICHR31_05	Continuation (6 of
18	141.5	4.2	2663	14 EBVBLR31	X99106 Epstein-Bar
19	140.5	4.2	8516	1 AE004995	AE004995 Pseudomon
20	140.5	4.2	11242	1 AE004995	AE004995 Halobacte
21	140.5	4.2	12041	1 AE010954	AE010954 Methanosa
22	140	4.1	161563	8 AP003076	AP003076 Oryza sat
23	138.5	4.1	139517	9 AC105446	AC105446 Homo sapi
24	138	4.1	4974	3 S83394	S83394 Insulin-lik
25	137	4.1	11482	1 AE001876	AE001876 Deinococc
26	136.5	4.0	27887	1 SCE50	AL163672 Streptomy
27	136	4.0	37620	1 SC1A6	AL023496 Streptomy
28	136	4.0	310050	1 RME603642	AL603642 Rhizobium
29	135	4.0	1891	1 SBSTRBF	X78972 S. bluenis
30	134.5	4.0	2535	9 AF152475	AF152475 Homo sapi
31	134.5	4.0	5254	9 AF152306	AF152306 Homo sapi
32	134.5	4.0	15742	1 AF013216	AF013216 Myxococcu
33	134.5	4.0	157970	9 AC005609	AC005609 Homo sapi
34	134	4.0	2661	14 AF172332	AF172332 Human her
35	133.5	4.0	10793	1 AE010748	AE010748 Methanosa
36	133.5	4.0	15274	1 AE004994	AE004994 Halobacte
37	132.5	3.9	1849	1 SGSTRB1F	X78973 S. galbus DS
38	132.5	3.9	10029	1 AE011825	AE011825 Xanthomon
39	132.5	3.9	11085	1 AE004450	AE004450 Pseudomon
40	132.5	3.9	25681	9 SAB414559	AB070950 Streptomy
41	132.5	3.9	117096	9 AC008468	AC008468 Homo sapi
42	131.5	3.9	3221	1 HALPHSOD	M2454 H. halobium
43	131.5	3.9	12730	1 AE005055	AE005055 Halobacte
44	131.5	3.9	21391	1 AB070950	AB070950 Streptomy
45	131	3.9	13164	1 AE005096	AE005096 Halobacte

RESULT 1

ALIGNMENTS

ARI25096
LOCUS ARI25096 2478 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 47 from patent US 6177075.
ACCESSION ARI25096
VERSION ARI25096.1 GI:14111158
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2478)
AUTHORS Christian,P.Daniel., Gordon,K.Hienrich,Julius. and
Hanzlik,T.Nelson.
TITLE Insect viruses and their uses in protecting plants
JOURNAL Patent: US 6177075-A 47 23-JAN-2001;
FEATURES
source Location/Qualifiers
BASP COUNT 516 a 854 c 607 g 501 t
ORIGIN /organism="unknown"

Alignment Scores:
Pred. No.: 3.15e-211 Length: 2478
Score: 3374.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-677-653a-50 (1-647) x ARI25096 (1-2478)

QY 1 MetGlyAspAlaGlyValAlaSerGlnArgProHisAsnArgGlyThrArgAsnVal 20
DB 366 ATGGGAATCTCTGGAGTGGCTCACAGCAGCTCACAACTGCGGACCCGTAACGTT 425
QY 21 ArgValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArgArgThrGly 40
DB 426 CCGGTGACGCCCAACACCCGTCACCGTCATGCTAGAAAGAAACCAACGCGCTGGACCGGA 485
QY 41 ArgGlnValSerProProAspAsnPheThrAlaAlaAlaGlnAspLeuAlaGlnSerLeu 60
DB 486 AGCCAGATTCTCCCTCGACAAATTTCACCGCTGCTGCACAAAGCTCGCGCAAGCTT 545
QY 61 AspAlaAsnThrValThrPheProAlaAsnIleSerSerMetProGluPheArgAsnTrp 80
DB 546 GAGCGCCAAACCGCTCACTTCCCGCTAACATCTAGCATGCGCGCAATTCCGGAAATTGG 605
QY 81 AlaIysGlyIleAspLeuAspSerAspSerIleGlyTrpTrpPheIleAsp 100
DB 606 GCCAAGGAAAGATCGACCTCGACCTCGATTCATCGGCTGGTGTACTTCAAGTACCTTGAC 665
QY 101 ProAlaGlyAlaThrGlnSerAlaArgAlaValGlyGluTrpSerIleProAspGly 120
DB 666 CCAGCGGGTGTCTACAGAGTCTGCGCGCGCTGCGGAGTACAGAAATCCCTGACGGC 725
QY 121 LeuValIlyPheSerValAspAlaGluIleArgGluIleTrpAsnGluGluCysProVal 140
DB 726 CTCTGAAGTCTCTCGCTCGACGAGATGAAGATCTATAACAGAGAGTCCCGCTC 785
QY 141 ValThrAspValSerValProLeuAspGlyArgGlnTrpSerLeuSerIlePheSerPhe 160
DB 786 GTACAGGACGTGCTCCCTCCCGCTCGACGGCGCGCACTGGAGCTCTCGATTTCTCTCTT 845
QY 161 PrometPheArgThrAlaTrpValAlaValAlaAsnValGluAsnIlysgLumetSerLeu 180
DB 846 CCGATGTTCAAGAACCCGCTACGTCGCGCGTAGCAAGCTGAGAAACAAGAGATGTCGCTC 905
QY 181 AspValValAsnAspLeuIleGluTrpLeuAsnAsnLeuAlaAspTrpArgTrpValVal 200
DB 906 GAGCTTGTCACGACCTCATCGAGTGGCTCAACAATCTGCGCCGACATGCGGCTTATGTCGTT 965
QY 201 AspSerGluGlnTrpIleAsnPheThrAsnAspThrThrTyrrTrpValArgIleArgVal 220
DB 966 GACTCTGAACAGTGTAACTTACCAATGACCAACGACGTAAGTCTCGCATCCGCGTT 1025

QY 221 LeuArgProThrTrpAspValProAspProThrGlnGluLeuValArgThrValSerAsp 240
DB 1026 CTACGTCCAACCTACGACAGCTTCCAGACCCACAGAGGCGCTTGTCCACAGTCTCGAC 1085
QY 241 TyrArgLeuThrTrpLysAlaIleThrCysGluAlaAsnMetProThrLeuValAspGln 260
DB 1086 TACCGCTCACTTATTAAGGGATGAATGTAACCAATGCAACATGCAACATCTGTCAGCAA 1145
QY 261 GlyPheTrpIleGlyGlnTrpAlaLeuThrProThrSerLeuProGlnTrpAspVal 280
DB 1146 GGGTTTGATCGGGCGCAGTACGCTCACCCCGCTAGCCGTACCGCAGTACGAGCTC 1205
QY 281 SerGlnAlaTrpAlaLeuHisThrLeuThrPheAlaArgProSerSerAlaAlaLeu 300
DB 1206 AGCAGGCTACAGCTCTGCGACACTTTTACCTTCCCGACACATCCAGCGCGCTGCATC 1265
QY 301 AlaPheValTrpAlaGlyLeuProGlnGlyGlyThrAlaProAlaGlyThrProAlaTrp 320
DB 1266 GCGTTTGTGGGCGAGGTTGGCCACAGGTGGCAGCTCGCCTGACAGCATCCAGCCTGG 1325
QY 321 GlnGlnAlaSerSerGlyGlyTrpLeuThrTrpArgHisAsnGlyThrThrPheProAla 340
DB 1326 GAGCAGGATCTCTGGGTGCTACTCTACCTGCGCCACAAAGGTACTCTTCCACGCT 1385
QY 341 GlySerValSerTrpValLeuProGlnGlyPheAlaLeuGluArgTrpAspProAsnAsp 360
DB 1386 GCGTCCCTTAGCTACGTTCTCCCTGAGAGGTTTGGCCTTACGCGCTACGACCCGGAAGAC 1445
QY 361 GlySerTrpThrAspPheAlaSerAlaGlyAspThrValThrPheArgGlnValAlaVal 380
DB 1446 GGGCTTTGGACCCACTTCCGTCGCGAGACACGCTCACTTCCGGCGAGTCCGCGCTC 1505
QY 381 AspGluValValValThrAsnAsnProAlaGlyGlySerAlaProThrPheThrVal 400
DB 1506 GACGAGGCTGTTGTACCAACAACCCCGCGGGGCGGCGCCCACTTACCGTG 1565
QY 401 ArgValProProSerAsnAlaTrpThrAsnThrValPheArgAsnThrLeuLeuGluThr 420
DB 1566 AGAGTGCGCCCTTCAAAAGGCTTACACCAACACCGTGTATGGAACAGCTTAGAGACT 1625
QY 421 ArgProSerSerArgArgLeuGluLeuProMetProProAlaAspPheGlyGlnThrVal 440
DB 1626 CGACCTCTCTCTGTAGGCTCGAATCCCTATCCACTGCTGACATTTGGACAGAGGCTC 1685
QY 441 AlaAsnAsnProLysIleGlnGlnSerLeuLeuGluThrLeuGlyCysTrpLeuVal 460
DB 1686 GCCAACAACCCGAAGATCGAGCTGCTTCTTAAAGAAACACTTGGCTGTATTGGCTC 1745
QY 461 HisSerLysMetArgAsnProValPheGlnLeuThrProAlaSerSerPheGlyAlaVal 480
DB 1746 CACTCCAAATATGGAAACCCCGTTTCCAGCTCACGCGACGAGCGTCTTGGCGCGTT 1805
QY 481 SerPheAsnAsnProGlyTrpGluArgThrArgAspLeuProAspTrpThrGlyIleArg 500
DB 1806 TCCTTCAACATCTCGGCTTATGAGCGCACACGCGGACCTCCGGAATCACTGTGCATCCG 1865
QY 501 AspSerPheAspGlnAsnMetSerThrAlaValAlaHisPheArgSerLeuSerHisSer 520
DB 1866 GACTCATTTGACCAACAATGTCACCGCTGTGGCCCACTTCCGCTCACTCTCCACTCC 1925
QY 521 CysSerIleValThrLysThrTrpGlnGlyTrpGluGlyValAlaThrAsnValAsnThrPro 540
DB 1926 TGCAGTATGCTGCTAAAGACTACAGGGGTGGGAAGGGGTACAGAAAGCTCAACAGCGCT 1985
QY 541 PheGlyGlnPheAlaHisAlaGlyLeuLeuLysAsnGluGluIleLeuCysLeuAlaAsp 560
DB 1986 TTCGGCCAAATTCGCGCAGCGGCGCTCTCAAGAAATGAGGAATCTCTGCTCGCGGAC 2045
QY 561 AspLeuAlaThrArgLeuThrGlyValTrpProAlaThrAspAsnPheAlaAlaVal 580
DB 2046 GACTGTGCAACCGCTGTACAGAGTGTCTACCCCGCACTGACAACATTTCCGGCGCGCTT 2105

QY 581 SerAlaPheAlaIalaAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIle 600
|||||
Db 2106 TCGCCCTCGCCGCGAAGATGCTGCTCCGTGTAAGTGGAGGCAACGCTCCATC 2165
QY 601 IleLysSerValGlyIuThrAlaValAlaAlaGlnSerGlyLeuAlaLysLeuPro 620
|||||
Db 2166 ATCAAGTCGCTGGCGAGACTGCGGCGGCTCAGTCAGTCCGCGGGAAGCTACCC 2225
QY 621 GlyLeuLeuMetSerValProGlyLysIleAlaAlaArgValArgAlaArgAlaArg 640
|||||
Db 2226 GGACTCTAATGATGTGTACAGGGAAGATGCGCGCTGTCGCGCGCGCGAGCGCGC 2285
QY 641 ArgArgAlaIalaArgAlaAsn 647
|||||
Db 2286 CGCCGCGCCCTGTCGCCAAT 2306
RESULT 2
ARI25097
LOCUS ARI25097 2478 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 49 from patent US 6177075.
ACCESSION ARI25097
VERSION ARI25097.1 GI:14111159
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2478)
AUTHORS Christian P. Daniel., Gordon, K. Heinrich, Julius. and
Hanzlik, T. Nelson.
TITLE Insect viruses and their uses in protecting plants
JOURNAL Patent: US 6177075-A 49 23-JAN-2001;
FEATURES Location/Qualifiers
source 1..2478
BASE COUNT 516 a 854 c 607 g 501 t
ORIGIN
Alignment Scores:
Pred. No.: 3,15e-211 Length: 2478
Score: 3374.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-677-653a-50 (1-647) x ARI25097 (1-2478)
QY 1 MetGlyAspAlaGlyValAlaSerGlnArgProHisAsnArgArgGlyThrArgAsnVal 20
|||||
Db 366 ATGGAGATGCTGGAGTGGGCTACAGCGACTCACACCGTCGCGGAACCGTAACGTT 425
QY 21 ArgValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArgArgThrGly 40
|||||
Db 426 CGGTTAGGCGCCAAACCGCTACCGCTCAATGATGAGAAACCAACGCGCTCGACCGGA 485
QY 41 ArgGlnValSerProProAspAsnPheThrAlaAlaAlaGlnAspLeuAlaGlnSerLeu 60
|||||
Db 486 AGGCAAGTTTCTCCCTGACATTTTCACCGCTGCTGCAACAGCTTCGCGCAAGCCTT 545
QY 61 AspAlaAsnThrValThrPheProAlaAsnIleSerSerMetProGluPheArgAsnTrp 80
|||||
Db 546 GAGCGCAACACGCTCACTTCCCGGCTAACATCTAGCATGCGCCGAATTCCGGAATTGG 605
QY 81 AlaLysGlyLysIleAspLeuAspSerAspSerIleGlyTrpTrpPheLysTrpLeuAsp 100
|||||
Db 606 GCCAAGGGAAGATCGACCTCGACTCGATTCGATCGGCTGATCTTCAAGTACCTTGAC 665
QY 101 ProAlaGlyAlaThrGlnSerAlaArgAlaValAlaGlyLysSerLysIleProAspGly 120
|||||
Db 666 CAGCGGGTCTACAGATCTGCGCGCGCGCTGCGGAGACTCGAAGATCCCTGACGGC 725
QY 121 LeuValLysPheSerValAspAlaGluIleArgGluIleTyrAsnGluGluCysProVal 140
|||||

Db 726 CTCGTCAAGTTCTCCGTCGACGAGATTAAGAGATCTATACGAGAGTCCCGCTC 785
QY 141 ValThrAspValSerValProLeuAspGlyArgGlnTrpSerLeuSerIlePheSerPhe 160
|||||
Db 786 GTCACTGACGTGTCGCTCCCGCTCGACGGCGGCGACAGGAGCTCTCGATTTCTCCTT 845
QY 161 PromePheArgThrAlaTyrValAlaValAlaAsnValGluLysLysGluMetSerLeu 180
|||||
Db 846 CCGATGTTCAAGAACCGCTACGTCGCGGTAGCGAAGCTCGAAGAACGAGATGTCCTC 905
QY 181 AspValValAsnAspLeuIleGluTrpLeuAsnLeuAlaAspTrpArgTrpValVal 200
|||||
Db 906 GAGCTTGTCAACGACCTCATCGAGTGGCTCAACAAATCTCCGCACTGCGTTATGTCCT 965
QY 201 AspSerGluGlnTrpIleAsnPheThrAsnAspThrThrTyrTyrValArgIleArgVal 220
|||||
Db 966 GACTCTGAACAGTGAATTACTTACCAATGACACACGCTACGTCGCGATCCGCTC 1025
QY 221 LeuArgProThrTyrAspValProAspProThrGluGlyLeuValArgThrValSerAsp 240
|||||
Db 1026 CTACGTCAACCTACGACGTTCCAGACCCACAGAGGCGCTTGTCCACAGTCTCAGAC 1085
QY 241 TyrArgLeuThrTyrLysAlaIleThrCysGluAlaAsnMetProThrLeuValAspGln 260
|||||
Db 1086 TACCGCTCACTTATAAGGGGATTAACATGTGAAGCCCAACATGCGACACTCGACCAA 1145
QY 261 GlyPheTrpIleGlyGlnTyrAlaLeuThrProThrSerLeuProGlnTyrAspVal 280
|||||
Db 1146 GCGTTTGGATCGGCGCGCGCTACCGCTCACCCTGAGCTTACCGCGAGTACGACGCTC 1205
QY 281 SerGluAlaThrAlaLeuHisThrLeuThrPheAlaArgProSerSerAlaAlaLeu 300
|||||
Db 1206 AGCAGGCTTACGCTGACACTTGTACCTTGCCAGACATCCAGCGCGCTGACCTC 1265
QY 301 AlaPheValTrpAlaGlyLeuProGlnGlyGlyThrAlaProAlaGlyThrProAlaTrp 320
|||||
Db 1266 GCGTTTGTGGGACAGGTTGGCCACAGGGTGGCAGCTGCGCTGACAGCATCCAGCTGG 1325
QY 321 GluGlnAlaSerSerGlyGlyTyrLeuThrTrpArgHisAsnGlyThrThrPheProAla 340
|||||
Db 1326 GACGAGGCAATCCGCGGTGCTACCTACCTGCGCGCACCAACGCTACTTCCAGACT 1385
QY 341 GlySerValSerTyrValLeuProGluGlyPheAlaLeuGluLysArgTyrAspProAsnAsp 360
|||||
Db 1386 GGCCTCGTTAGCTTACGTTTCCCTGAGGGTTTGGCCCTTACGGGCTTACGACCCGAAGAC 1445
QY 361 GlySerTrpThrAspPheAlaSerAlaGlyAspThrValThrPheArgGlnValAlaVal 380
|||||
Db 1446 GGCTCTTGACCGCACTTCGCTCCGAGAGACACCGTCACTTCCGGCAGTCCGCGCTC 1505
QY 381 AspGluValValValThrAsnAsnProAlaGlyGlyGlySerAlaProThrPheThrVal 400
|||||
Db 1506 GACGAGGTCGTTGACCAACACCCCGCGGGGCGGCGCCGCCACCTTACCGCTG 1565
QY 401 ArgValProProSerAsnAlaTyrThrAsnThrValPheArgAsnThrLeuLeuGluThr 420
|||||
Db 1566 AGAGTGCGCCCTTCAACCGCTTACACCAACACCGTGTATTGGAACACGCTTAGAGACT 1625
QY 421 ArgProSerSerArgArgLeuGluLeuProMetProProAlaAspPheGlyGlnThrVal 440
|||||
Db 1626 CGACCTCTCTCTGCTGAGGCTCGAATCCCTATGCACTGCTGATTTGGACAGAGGCTC 1685
QY 441 AlaAsnAsnProLysIleGluGlnSerLeuLeuLysGluThrLeuGlyCysTyrLeuVal 460
|||||
Db 1686 GCCAACCAACCGAAGATCGAGTGCCTTCTTAAGAAACACTTGGCTGCTATTGGCTC 1745
QY 461 HisSerLysMetArgAsnProValPheGlnLeuThrProAlaSerSerPheGlyValVal 480
|||||
Db 1746 CACTTCAAAATGCGAAACCCGTTTTCACACTACGCGCAACGACTCTTTGGCGCGCTT 1805
QY 481 SerPheAsnAsnProGlyTyrGluArgThrArgAspLeuProAspTyrThrGlyIleArg 500
|||||
Db 1806 TCCTTCAACAATCCGGGTTATGAGCGCACGCGGACTCTCCGGAGTACACTGCAATCCGT 1865

QY 501 AspSerPheAspGlnAsnMetSerThrAlaValAlaHisPheArgSerLeuSerHisSer 520
| | | | |
Db 1866 GACTCATTCGACACAGACATGTCACACCGCTGTGGCCCACTTCGCCGCTATCTCCACCTCC 1925
QY 521 CysSerIleValThrLysThrTyrglnGlyTyrPglGlyValThrAsnValAsnthrPro 540
| | | | |
Db 1926 TTCACATATCGTCACTAAGACCTACACGAGGTGGGAAGCGGTACAGAACCTCAACAGCGCT 1985
QY 541 PheGlyGlnPheAlaHisAlaGlyLeuLeuLysAsnGluGlnIleLeuGlyLeuAlaAsp 560
| | | | |
Db 1986 TTCGGCCATTCGCGACCGGCGGCTCTCAAGAAATGAGAGATCCTCTGCTCGCGGAC 2045
QY 561 AspLeuAlaThrArgLeuThrGlyValTyrProAlaThrAspAsnPheAlaAlaVal 580
| | | | |
Db 2046 GACCTGGCCACCGCTGTCAAGAGTGTCTACCCGCCGACAGACACTTCGCCGCGCGCTT 2105
QY 581 SerAlaPheAlaAlaAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIle 600
| | | | |
Db 2106 TCTGCTTCGCGCGCAACATGCTGTCTCTCCGTGTAAGTCGAGGCAACGCTCTCCATC 2165
QY 601 IleLysSerValGlyGlnThrAlaValGlyAlaAlaGlnSerGlyLeuAlaLysLeuPro 620
| | | | |
Db 2166 ATCAAGTCCGTTGGGAGAGCTCCGTCGCGCGCTCAGTCCGCGCTCGGAAAGTACCC 2225
QY 621 GlyLeuLeuMetSerValProGlyLysIleAlaAlaArgValArgAlaArgAlaArg 640
| | | | |
Db 2226 GGACTGCTTAATGATGATACCAAGAAAGATTGCCGCGCTGTCGCGCGCGCGAGCGCC 2285
QY 641 ArgAlaGlnAlaAlaArgAlaAsn 647
| | | | |
Db 2286 CCGCGCGCGCTGTGCTGCTT 2306
RESULT 3
ARI59968
LOCUS ARI59968 2478 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6251654.
ACCESSION ARI59968
VERSION ARI59968.1 GI:16222852
KEYWORDS
SOURCE .
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2478)
AUTHORS Gordon,K.,Heinrich, and Hanzlik,T. Nelson.
TITLE Modified small RNA viruses
JOURNAL Patent: US 6251654-A 1 26-JUN-2001;
FEATURES
Location/Qualifiers
1..2478
Source
BASE COUNT 516 a 854 c 607 g 501 t
ORIGIN
Alignment Scores:
Pred. No.: 3.15e-211 Length: 2478
Score: 3374.00 Matches: 647
Percent Similarity: 100.00% Conservatave: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-677-653A-50 (1-647) x ARI59968 (1-2478)
QY 1 MetGlyAspAlaGlyValAlaLaserGlnArgProHisAsnArgArgGlyThrArgAsnVal 20
| | | | |
Db 366 ATGGAGATGCTGGATGGCGCTACAGCGACCTCAACACCGTCGCGGAACCGCTAACGTT 425
QY 21 ArgValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArgArgArgThrGly 40
| | | | |
Db 426 CGGGTCACCGCAACACCGCTACCGCTCAATGTTAGAAACCAACGCGCTCGGAGCGGA 485
QY 41 ArgGlnValSerProProSerAsnPheThrAlaAlaGlnAspLeuAlaGlnSerLeu 60
| | | | |

Db 486 AGGCAAGTTTCTCCCTGACAAATTCACCGCTGCTGCACAAAGACCTCGCGCAAAAGCTT 545
QY 61 AspAlaAsnThrValThrPheProAlaAsnIleSerSerMetProGluPheArgAsnthrPro 80
| | | | |
Db 546 GAGGCCAACACGCTGACTTCCCGCTAACATCTCAGATGCCGAAATTCGGAAATTTG 605
QY 81 AlaLysGlyLysIleAspLeuAspSerAspSerIleGlyTyrPheLysThrLeuAsp 100
| | | | |
Db 606 GCCAAGGGAAGATGACATCGACTCCGATTCATCGGCTGTTACTTCAAAGTACCTGTGAC 665
QY 101 ProAlaGlyAlaThrGlnSerAlaArgAlaValAlaGlyLysThrLysIleProAspGly 120
| | | | |
Db 666 CCAGCGGGGTCTACAGAGCTCGCGCGCTCGGCGAGTACTCCAAAGTTCCCTCAACGCG 725
QY 121 LeuValLysPheSerValAspAlaGluIleArgGluIleTyrAsnGluGlnLysProVal 140
| | | | |
Db 726 CTCGTCAACTTCTCCGTCACGCAGAGATTAAGAGATTTATACGAGAGATGCCCCGTC 785
QY 141 ValThrAspValSerValProLeuAspGlyArgGlnTyrSerLeuSerIlePheSerPhe 160
| | | | |
Db 786 GTCACTGACGTGTCGCTCCCTCGACGCGCGCCAGTGAGGCTCTCGATTTCCTCTT 845
QY 161 PrometPheArgThrAlaTyrValAlaValAlaAsnValGluAsnLysGluMetSerLeu 180
| | | | |
Db 846 CCGATGTTTCAGAACCGCTACCTCGCCGCTGAGCAACGTCGAAACAGAGAGATGTCGCTC 905
QY 181 AspValAlaAsnAspLeuIleGlnThrPleuAsnAsnLeuAlaAspThrArgTyrAlaVal 200
| | | | |
Db 906 GACGTTGTACAGACTCTCATGAGTGGCTCAACAAATCTCGCGACTGGGTTATGCTT 965
QY 201 AspSerGluGlnTyrIleAsnPheThrAsnAspThrThrTyrTyrValArgIleArgVal 220
| | | | |
Db 966 GACTCTGAACAGTGTGATTAATCAATCAACATGACACAGCTACTAGTCGCGATTCGCGCTT 1025
QY 221 LeuArgProThrTyrAspValProAspProThrGlnGlyLeuValArgThrValSerAsp 240
| | | | |
Db 1026 CTACGTCCAACCTACGACCTCCACACCCACAGAGGCGCTTGTTCGACACTCTCAGAC 1085
QY 241 TyrArgLeuThrTyrLysAlaIleThrCysGluAlaAsnMetProThrLeuValAspGln 260
| | | | |
Db 1086 TACCGCTTCACCTTAATAGGCGGTAACATGTGAAGGCAACATGCCAATCTGTCACCA 1145
QY 261 GlyPheThrIleGlyGlyGlnTyrAlaLeuThrProThrSerLeuProGlnTyrAspVal 280
| | | | |
Db 1146 GCGTTTGGATCGGGGCGCAGTACCTCTCACCCGACAGCTACCGAGTACAGACGTC 1205
QY 281 SerGluAlaTyrAlaLeuHisThrLeuThrPheAlaArgProSerSerAlaAlaLeu 300
| | | | |
Db 1206 AGCGAGGCTTACGCTCTGACACTTGTGACCTTCGCCAGACCATCCAGCGCCCTGCACTC 1265
QY 301 AlaPheValThrAlaGluLeuProGlnGlyThrAlaProAlaGlyThrProAlaThr 320
| | | | |
Db 1266 GCGTTGTGTGGGCAAGTTTGCACAGGTGGGCTGCGCTGCAAGCACTCCACCTG 1325
QY 321 GlnGlnAlaSerSerGlyGlyTyrLeuThrThrPACGHisAsnGlyThrThrPheProAla 340
| | | | |
Db 1326 GAGCAGGCGATCTCGGGTGGCTTCACTCACTGCGCCCAACAGTACTTCCACACT 1385
QY 341 GlySerValSerTyrValLeuProGlnGlyPheAlaLeuGluArgTyrAspProAsnAsp 360
| | | | |
Db 1386 GGCTCCGTTAGCTACGTTCTCCCTGAGGGTTTCGCCCTTGAGCGCTACGACCCGAACGAC 1445
QY 361 GlySerTyrPheThrAspPheAlaSerAlaGlyAspThrValThrPheArgGlnValAlaVal 380
| | | | |
Db 1446 GGCTCTTGAGCCGACTTGTCTCCCGCAGAGACACGCTCACTTCCGAGGTCCCGCTC 1505
QY 381 AspGluValAlaValThrAsnAsnProAlaGlyGlyGlySerAlaProThrPheThrVal 400
| | | | |
Db 1506 GACGAGGTGCTGTGACCAACACCCCGCGCGCGGCGAGCGCCCACTTACCGCTG 1565
QY 401 ArgValProProSerAsnAlaTyrThrAsnThrValPheArgAsnThrLeuLeuGlnThr 420
| | | | |
Db 1566 AGAGTGGCCCTTCAACAGCTTACACCAACACCGTGTTAGAAGACGCTTAGAGACT 1625

QY 421 ArpProserSerArgArgLeuGluLeuPrometProProAlaaspheGlyGlnThrVal 440
|||||
Db 1626 CGACCTCTCTCTGAGGCTGAACTCCCTATGCGACCTGCTGACTTGGACAGCGGCTC 1685
QY 441 AlaAsnAsnProLysIleGluGlnSerLeuLeuLysGlnThrIleGlyCysTyrLeuVal 460
|||||
Db 1686 GCCAACAACCCGAGAGTCGACACTCGCTTAAAGAAACACTTGGCTGATTTGGTTC 1745
QY 461 HisSerLysMetArgAsnProValPheGlnLeuThrProAlaSerSerPheGlyAlaVal 480
|||||
Db 1746 CACCTCAAAATGCGAAACCCCGTTTCCAGCTCACGCCAGCAGCTCTCTTGGCGCCCTT 1805
QY 481 SerPheAsnAsnProGlyTyrGluArgThrArgAsnLeuProAspTyrThrGlyIleArg 500
|||||
Db 1806 TCCTTCACAACTCCGGGTTATGAGCGCACAGCGACCTCCGAGCTACACTGGCATCCGT 1865
QY 501 AspSerPheAspGlnAsnMetSerThrAlaValAlaHisPheArgSerLeuSerHisSer 520
|||||
Db 1866 GACTCATTCGACGAGAACATGTCACCGCTGGCCACTTCCGCTCATCTCCACATCC 1925
QY 521 CysSerIleValIleThrLysThrTyrGlnGlyTyrPheGlyValIleThrAsnValAsnThrPro 540
|||||
Db 1926 TGCAGTATCCGCTACAGACCTACAGGCTTGGGAGGCGCTCAGCAACGTCACACGCT 1985
QY 541 PheGlyGlnPheAlaHisAlaGlyLeuLeuLysAsnGlnGluIleLeuCysLeuAlaAsp 560
|||||
Db 1986 TTCGGCCAAATTCGCGACGCGGCGCTCTCTCAAGATGAGAGATCTCTGCTCGCCGAC 2045
QY 561 AspLeuAlaThrArgLeuThrGlyValTyrProAlaThrAspAsnPheAlaAlaVal 580
|||||
Db 2046 GACCTGGCCACCGCTCTCAGAGTGTCTACCCCGCACTGACAACTTGGCGCGCCCTT 2105
QY 581 SerAlaPheAlaAlaAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIle 600
|||||
Db 2106 TCTGCTTCGCGCGGAAACATCTGCTCCCTGCTGAAGTGGAGGCAACGCTCCCTCAC 2165
QY 601 IleLysSerValGlyGlnThrAlaValGlyAlaGlnSerGlyLeuAlaIleLeuPro 620
|||||
Db 2166 ATCAAGTCCTGTTGGCGAGACTGCGTGGCGCGGCTCAGTCCGGCTCGCGAAGCTACCC 2225
QY 621 GlyLeuLeuMetSerValProGlyLysIleAlaAlaArgValArgAlaArgAlaArg 640
|||||
Db 2226 GGACTGGTAATGAGTGTACAGGGAAGATTCGCGGCGGTGTCCGGCGCGCGAGCGGCG 2285
QY 641 ArgArgAlaAlaArgAlaAsn 647
|||||
Db 2286 CGCGCGCGCGCTCGTGCCTAAT 2306
RESULT 4
AX028950
LOCUS AX028950 2478 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 1 from Patent WO974666.
ACCESSION AX028950
VERSION AX028950.1 GI:10189993
KEYWORDS
SOURCE Helicoverpa armigera stunt virus.
ORGANISM Helicoverpa armigera stunt virus
Virus; ssRNA positive-strand viruses, no DNA stage; Tetraviridae;
unclassified Tetraviridae.
1 (bases 1 to 2478)
REFERENCE
AUTHORS Gordon, K.H. and Hanzlik, T.N.
TITLE Modified small rna viruses
JOURNAL Patent: WO 974666-A 1 11-DEC-1997;
COMM SCIENT IND RES ORG (AU) ; GORDON KARL HEINRICH (AU) ; HANZLIK
TERRY NELSON (AU)
FEATURES
source location/qualifiers
1..2478
/organism="Helicoverpa armigera stunt virus"
/db_xref="taxon:37206"
BASE COUNT 516 a 854 c 607 g 501 t

Alignment Scores:
Pred. NO.: 3,15e-211 Length: 2478
Score: 3374.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-677-653a-50 (1-647) x AX028950 (1-2478)
QY 1 MetGlyAspAlaGlyValAlaSerGlnArgProHisAsnArgArgGlyThrArgAsnVal 20
|||||
Db 366 ATGGAGATGCTGGAGTGGCTGACAGAGGACTCACAACCGTCCGGGAACCGTAAGCTT 425
QY 21 ArgValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArgArgThrGly 40
|||||
Db 426 CGGGTCACGGCCCAACACCGTCAATGATGAGAAACCAACACGCGCTCGGACCGGA 485
QY 41 ArgGlnValSerProProAspAsnPheThrAlaAlaGlnAspLeuAlaGlnSerLeu 60
|||||
Db 486 AGCGAAGTTTCTCCCTGACAAATTTCCAGCTGCTGACAGACCTCGCCAAAGCTT 545
QY 61 AspAlaAsnThrValThrPheProAlaAsnIleSerSerMetProGluPheArgAsnThr 80
|||||
Db 546 GACGCGCAACACCGTCACCTTCCCGCTAACATCTCTACATGCCGGAATTCGGAAATGG 605
QY 81 AlaLysGlyLysIleAspLeuAspSerAspSerIleGlyTyrPtyrPheLysTyrLeuAsp 100
|||||
Db 606 GCCAAGGGAAGATCGACCTGACCTCCATTCATCCGCTGCTACTTCMACTACCTTGAC 665
QY 101 ProAlaGlyAlaThrGlnSerAlaArgAlaValGlyGluTyrSerLysIleProAspGly 120
|||||
Db 666 CCAGCGGCTGTACAGATCTGCGCGCGCGCTGGCGAGTACTGAAAGATCTCTGACGCG 725
QY 121 LeuValLysPheSerValAspAlaGluIleArgGluIleTyrAsnGluGluCysProVal 140
|||||
Db 726 CTCGTCAGATCTCCGCTGCGAGCAGAGATTAAGAGATCTTAACGAGAGCTGCCCGCTC 785
QY 141 ValThrAspValSerValProLeuAspGlyArgGlnTyrPserLeuSerIlePheSerPhe 160
|||||
Db 786 GTCACTGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 845
QY 161 PrometPheArgThrAlaTyrValAlaValAlaAsnValGluAsnLysGluMetSerLeu 180
|||||
Db 846 CCGATGTCAGACCGCTACGTCGCTGACGAGACGTCGAGAACAGAGATGTCGCTC 905
QY 181 AspValValaAsnAspLeuIleGluThrLeuAsnAsnLeuAlaAspTyrArgTyrVal 200
|||||
Db 906 GACGTCGCAACGACCTCATGAGTGGCTCAACAATCTCGCGACTGGCGTTATGTGCTT 965
QY 201 AspSerGluGlnTyrIleAsnPheThrAsnAspThrTyrTyrValArgIleArgVal 220
|||||
Db 966 GACTCTGAACAGTGGATTAACCTTCAACAATGACACACAGTACTGTCGCAATCGGCTT 1025
QY 221 LeuArgProThrTyrAspValProAspProThrGluGlyLeuValArgThrValSerAsp 240
|||||
Db 1026 CTAGCTCAACCTACGAGCTTCCACAGCCACAGAGGCGCTTGTTCGCAAGTCTCAAC 1085
QY 241 TyrArgLeuThrTyrLysAlaIleThrCysGluAlaAsnMetProThrLeuValAspGln 260
|||||
Db 1086 TACCGCTCACTTAAAGGCGATTAACATGTAAGCGCAACATGCCAACACTGTCGACCAA 1145
QY 261 GlyPheThrIleGlyGlyGlnTyrAlaLeuThrProThrSerLeuProGlnTyrAspVal 280
|||||
Db 1146 GCGTTTGTGATCGGGCGGCGCACTACGCTTCAACCCGACTACCTAACCCAGTACGACGTC 1205
QY 281 SerGluAlaTyrAlaLeuHisThrLeuThrPheAlaArgProSerSerAlaAlaAlaLeu 300
|||||
Db 1206 AGCGAGGCTACGCTGTGCAACATTTGACCTTCCGCAACCATTCAGCGCGCTGACATC 1265
QY 301 AlaPheValTyrAlaGlyLeuProGlnGlyGlyThrAlaProAlaGlyThrProAlaThr 320
|||||
Db 1266 GCGTTTGTGTGGGCAAGTTTGCACAGGTTGGCACCTGCGCTGCAAGGCACTCCAGCTTGG 1325

```

QY 321 GUGLAlalaserSerGlyGlyTyrLeuThrTrpArgHisAsnGlyThrThrPheProAla 340
    |||||||
Db 1326 GACAGAGCATCTCGGGTGGCTACTACTCGGGCCACAAAGCTACTCTCCAGCT 1385
QY 341 GlySerValSerTyrValLeuProGluGlyPheAlaLeuGluArgTyrAspProAsnAsp 360
    |||||||
Db 1386 GGCTCGGTTAGTACGTTTCCCTGAGAGGGTTGGCCCTTGAGGCGTACGACCCGAAGGAC 1445
QY 361 GlySerThrPheAspPheAlaSerAlaGlyAspThrValThrPheArgGlnValAlaVal 380
    |||||||
Db 1446 GGCTCTGGACCGACTTCGCTCCGACGAGACCGCTCACTTCCGGCAGTCCGCGTC 1505
QY 381 AspGluValValValThrAsnAsnProAlaGlyGlySerAlaProThrPheThrVal 400
    |||||||
Db 1506 GACGAGTGTCTTGACCAACAAACCCCGCGGGCGGACGCCGCCCTCACCCGTG 1565
QY 401 ArgValProProSerAsnAlaTyrThrAsnThrValPheArgAsnThrLeuLeuGluThr 420
    |||||||
Db 1566 AGAGTCCCCCTTCAAAACGCTTACACCAACACCGGTGTTAGGAACAGCTCTTAGAGACT 1625
QY 421 ArgProSerSerArgArgLeuGluLeuProMetProProAlaAspPheGlyGlnThrVal 440
    |||||||
Db 1626 CGACCTCTCTCTGAGGCTCGAAGTCCCTATGCCACCTGCTGACCTTGGACAGAGCGTC 1685
QY 441 AlaAsnAsnProLysIleGluGlnSerLeuLeuGlyLutThrLeuGlyCysTyrLeuVal 460
    |||||||
Db 1686 GCCAACACCCGGAAGATCGAGCAGCTCGCTTAAAGAAACACTTGGCTCTATTGGTC 1745
QY 461 HisSerLysMetArgAsnProValPheGlnLeuThrProAlaSerSerPheGlyAlaVal 480
    |||||||
Db 1746 CACTTCAAAATGCGAAACCCCGTTTCCACCTCACGCCACCGACCTTGGCGCGCTT 1805
QY 481 SerPheAsnAsnProGlyTyrGluArgThrArgAspLeuProAspTyrThrGlyIleArg 500
    |||||||
Db 1806 TCCTTAAACATCCGGGTTATGAGCGACACGGGACCTCCGGAGTACATGGCATCCGT 1865
QY 501 AspSerPheAspGlnAsnMetSerThrAlaValAlaHisPheArgSerLeuSerHisSer 520
    |||||||
Db 1866 GACTCATTTGACCAAGATGTCACCGGTGGCGGACACTTCCGCTCCTCCACTCC 1925
QY 521 CysSerIleValThrLysThrTyrGlnGlyTyrGluGlyValThrAsnValAsnThrPro 540
    |||||||
Db 1926 TGCAGTATGCTCACTAAGACTCTACAGGGTGGGAAGGCTCTCAAGACGCAACAGCCCT 1985
QY 541 PheGlyGlnPheAlaHisAlaGlyLeuLeuLysAsnGluLutLeuGlySerLeuAlaAsp 560
    |||||||
Db 1986 TTCGGCAATTCGGCAGCGGGGCTCTCAAGAAATGAGAGATCCCTGCTCCGCGAC 2045
QY 561 AspLeuAlaThrArgLeuThrGlyValTyrProAlaThrAspAsnPheAlaAlaVal 580
    |||||||
Db 2046 GACCTGGCCACCGCTGCACAGGTGTCTACCCCGCCACTACAACTTCCGGGCGCGCTT 2105
QY 581 SerAlaPheAlaAlaAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIle 600
    |||||||
Db 2106 TCTGCTTGGCCCGCAACATGCTGCTCTCGCTCAAGTCGGAAGCAACGCTCTCATC 2165
QY 601 IleLysSerValGlyLutThrAlaValAlaGlnSerGlyLeuAlaLysLeuPro 620
    |||||||
Db 2166 ATCAATTCGTTGGCAGACTGCGTGGGGGCTCAAGTCGCGCCCTCGGGAAGCTACCC 2225
QY 621 GlyLeuLeuMetSerValProGlyLysIleAlaAlaArgValAlaArgAlaArgAlaArg 640
    |||||||
Db 2226 GGACTGCTATAGTGTACAGGGAAGATTCCGGCGGTGTCCGCGCGCGGAGGGCGC 2285
QY 641 ArgArgAlaAlaArgAlaAsn 647
    |||||||
Db 2286 CGCCGCGCGCTGTGCAAT 2306

```

```

RESULT 5
AX028971/c AX028971 2478 bp DNA linear PART 16-SEP-2000
LOCUS Sequence 22 from Patent WO9746666.
DEFINITION

```

```

ACCESSION AX028971 GI:10190009
VERSION AX028971.1
KEYWORDS
SOURCE Helicoverpa armigera stunt virus.
ORGANISM Helicoverpa armigera stunt virus
REFERENCE 1 (bases 1 to 2478)
AUTHORS Gordon, K.H. and Hanzlik, T.N.
TITLE Modified small rna viruses
JOURNAL Patent: WO 9746666-A 22 11-DEC-1997;
TERRY NELSON (AU)
FEATURES
    source location/Qualifiers
        1..2478 /organism="Helicoverpa armigera stunt virus"
        /db_xref="taxon:37206"
BASE COUNT 501 a 607 c 854 g 516 t
ORIGIN
Alignment Scores:
Pred. No.: 3,15e-211 Length: 2478
Score: 3374.00 Matches: 647
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-677-653a-50 (1-647) x AX028971 (1-2478)
QY 1 MetGlyAspAlaGlyValAlaSerGlnArgProHisAsnArgGlyThrArgAsnVal 20
    |||||||
Db 2113 ATGGAGATGCTGAGAGGGGCTACAGCGACTCACACCGCTGGCGGAACCCGTAACGTT 2054
QY 21 ArgValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArgArgGly 40
    |||||||
Db 2053 CGGGTCAGCCCAACACCGCTCACCGTCMAAGTGAAGAAACCAACGGCGTGGACCGGA 1994
QY 41 ArgGlnValSerProProAspAsnPheThrAlaAlaGlnAspLeuAlaGlnSerLeu 60
    |||||||
Db 1993 AGCCAAATTTCTCCCTCCCTGACATTTTCACCGCTGCTGACACAAACCTCGGCAACGCTT 1934
QY 61 AspAlaAsnThrValThrPheProAlaAsnIleSerSerMetProGluPheArgAsnThrPro 80
    |||||||
Db 1933 GACGCCAACACCGTCACTTCCCGCTACATCTTACGATGCGCCGAATTCGGAATTTGG 1874
QY 81 AlaLysGlyLysIleAspLeuAspSerSerIleGlyTyrPheLysTyrLeuAsp 100
    |||||||
Db 1873 GCCAAGGGAAGATCGACCTCGACTCCGATTCATGGGTGATCTTCAAGTACCTTGAC 1814
QY 101 ProAlaGlyAlaThrGlnSerAlaArgAlaValAlaGlyLutTyrSerLysIleProAspGly 120
    |||||||
Db 1813 CCAGCGGGTCTACAGAGTCTGCGCGCGCGCTGGCGAGTACTCGAAGATCCCTGACGGC 1754
QY 121 LeuValLysPheSerValAspAlaGluIleArgGluIleTyrAsnGluGluCysProVal 140
    |||||||
Db 1753 CTCGTCAAGTTTCCGCGACGCGAGATTAAGAGATCTATTAACGAGAGTCCCGCTTC 1694
QY 141 ValThrAspValSerValProLeuAspGlyArgGlnTyrSerLeuSerIlePheSerPhe 160
    |||||||
Db 1693 GTCACTGAGCTGTCCCTCCCTCGACGGCGCGCAATGGAGCTCTCGATTTTCTCTTTT 1634
QY 161 ProMetPheArgThrAlaTyrValAlaValAlaAsnValAlaGlnLysAsnGluMetSerLeu 180
    |||||||
Db 1633 CCGATGTTCAAGAACCGGCTTACGCTCGCGCTGACAGCTGAGAAACAAGAGATGTGCTC 1574
QY 181 AspValValAsnAspLeuIleGluTyrPheAsnAsnLeuAlaAspThrArgTyrValVal 200
    |||||||
Db 1573 GACGTTGTCAAGCACTCATCGAGTGGCTCAACAATCTGCGCAGTGGGTTATGTGCTT 1514
QY 201 AspSerGluGlnThrPheAsnPheThrAsnAspThrThrTyrTyrValArgIleArgVal 220
    |||||||
Db 1513 GACTCTGAACAGTGGATTAATTACCAATGACACACGTAACACTGCTCGCATCCCGCTT 1454

```

QY 221 LeuArgProThrTyrAspValProAspProThrGluGluValArgThrValSerAsp 240
 |||||
 Db 1453 CTACGTCACACCTACGACGCTCCAGACCCACAGAGGCGCTTGTCCACAGTCTCAGAC 1394
 QY 241 TyrArgLeuThrTyrLysAlaIleThrCysGluAlaAsnMetProThrLeuValAspGln 260
 |||||
 Db 1393 TACCGCTCACTTATAAGGCGGATACATGTGAAGCCAAACATGGCAACACTCGTCGACCAA 1334
 QY 261 GlyPheTrpIleGluGluGluTyrAlaLeuThrProThrSerLeuProGlnTyrAspVal 280
 |||||
 Db 1333 GGGTTTGGATCGCGCGCCAGTACGCTCACCCCGACTGACCTACCGCAGTACGAGCTC 1274
 QY 281 SerGlnAlaTyrAlaLeuHisThrIleThrPheAlaArgProSerSerAlaAlaLeu 300
 |||||
 Db 1273 AGCGAGGCTACGCTGCGACCTTGTACCTTGGCCAGACACTCCAGCGCGCTGCGACTC 1214
 QY 301 AlaPheValTrpAlaGluLeuProGlnGluGlyThrAlaProAlaGlyThrProAlaTrp 320
 |||||
 Db 1213 GCGTTGTGTGGGAGGTTTGGCCACAGGCGGACCTCGGCTCGAGGACTCCAGCCTGG 1154
 QY 321 GluGlnAlaSerSerGlyGlyTyrLeuThrTrpArgHisAsnGlyThrPheProAla 340
 |||||
 Db 1153 GACGAGGATCTCGGGGTGGCTACTCAGCTGCGCCACCAACGGTACTTCTCCAGCT 1094
 QY 341 GlySerValSerTyrValLeuProGluGluPheAlaLeuGluArgTyrAspProAsnAsp 360
 |||||
 Db 1093 GGGTCGGTAAAGTACGCTTCCCTCGAGAGGTTTGGCCCTTGAAGCGTACGACCGCAAGAC 1034
 QY 361 GlySerTrpThrAspPheAlaSerAlaGlyAspThrValThrPheArgGlnAlaVal 380
 |||||
 Db 1033 GGGCTTTGGACGACTTCGCTCCGACGAGACACCTCACTTCCGGCAGTCCGCGTC 974
 QY 381 AspGluValValThrAsnAsnProAlaGlyGlySerAlaProThrPheThrVal 400
 |||||
 Db 973 GACGAGGCTGTTGACCAACAACCGCGCGGCGGAGCGCCCACTTCCAGCGTG 914
 QY 401 ArgValProProSerAsnAlaTyrThrAsnThrValPheArgAsnThrLeuGluThr 420
 |||||
 Db 913 AGAGTGGCCCTTCAAGCGTTTACCAACACCGTGTGTGAAGAACGCTTGAAGACT 854
 QY 421 ArgProSerSerArgArgLeuGluLeuProMetProProAlaAspPheGlyGlnThrVal 440
 |||||
 Db 853 CGACCCCTCTCTGTAGGCTCGAACTCCTATGCACTCTGACTTGGACAGACGCTC 794
 QY 441 AlaAsnAsnProLysIleGluGlnSerLeuLeuLysGluThrLeuGlyCysTyrLeuVal 460
 |||||
 Db 793 GCCAACAAACCGAAGATCGAGAGTGGCTTCTTAAAGAAACACTTGGCTGCTATTGGTC 734
 QY 461 HisSerLysMetArgAsnProValPheGlnLeuThrProAlaSerSerPheGlyAlaVal 480
 |||||
 Db 733 CACTCCAAATGGGAAACCCCGTTTTCACCTCAGCGCACGCTCTTGGGCGGCTT 674
 QY 481 SerPheAsnAsnProGlyTyrGluArgThrArgAspLeuProAspTyrThrGlyIleArg 500
 |||||
 Db 673 TCTTTCACACATCGGGTTTATGAGCGACACGAGACTCCCGGACTACACTGGCATCGT 614
 QY 501 AspSerPheAspGlnAsnMetSerThrAlaValAlaHisPheArgSerLeuSerHisSer 520
 |||||
 Db 613 GACTCATTCGACCAAGACATGTCACGCGTGTGGCCCACTTCGCTCAGCTCCCACTCC 554
 QY 521 CysSerIleValThrLysThrTyrGlnGlyTyrGluGlyValThrAsnValAsnThrPro 540
 |||||
 Db 553 TGGAGATCTCTCACTAAGACTCAGAGGTTGGGAAAGCGCTCAAGAACACACGCGCT 494
 QY 541 PheGlyGlnPheAlaHisAlaGlyLeuLeuLysAsnGluGluIleLeuGlyLeuAlaAsp 560
 |||||
 Db 493 TTGGGCAATTCGGCGACGGGGGCTCTCAAGATATGAGATGCTCTGCTCGCCGAC 434
 QY 561 AspLeuAlaThrArgLeuThrGlyValTyrProAlaThrAspAsnPheAlaAlaVal 580
 |||||
 Db 433 GACCTGGCCACCGGTCTCAGAGGTGTACCCGCGCACTGACAACTTCGCGCGCGCTT 374

QY 581 SerAlaPheAlaAlaAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIle 600
 |||||
 Db 373 TCTGCTTCCCGCGGAACATGCTCTCTCCGTCTGAAGTCGAGGACGACTCTCCATC 314
 QY 601 IleLysSerValGlyGluThrAlaValAlaAlaAlaGlnSerGlyLeuAlaLysLeuPro 620
 |||||
 Db 313 ATCAAGTCGCTGGCGAGACTGCGGTGGCGGCGCTCAGTCGCGCTCGGAAGCTACCC 254
 QY 621 GlyLeuLeuMetSerValProGlyLysIleAlaAlaArgValAlaArgAlaArg 640
 |||||
 Db 253 GGACTGCTAATGAGTGTACCGAGGAAATTCGCGCGTGTCCGCGCGCGAGCGCGC 194
 QY 641 ArgArgAlaAlaArgAlaAsn 647
 |||||
 Db 193 CGCCGCGCGCTCTGCTCCAAAT 173
 RESULT 6
 HVScoat
 LOCUS 2478 bp ss-RNA linear VRL 11-MAY-1995
 DEFINITION Helicoverpa armigera stunt virus RNA2 coat protein (p71) and p17
 ACCESSION U37299.1 GI:609503
 VERSION 1.37299
 KEYWORDS coat protein; transfer RNA-like structure.
 SOURCE Helicoverpa armigera stunt virus (individual isolate black mountain) cDNA to genomic RNA.
 ORGANISM Helicoverpa armigera stunt virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Tetraviridae; unclassified Tetraviridae.
 1 (bases 1 to 2478)
 Hanzlik, T.N., Dorrian, S.J., Johnson, K.N., Brooks, E.M. and Gordon, K.H.
 TITLE Sequence of RNA2 of the Helicoverpa armigera stunt virus (Tetraviridae) and bacterial expression of its genes
 JOURNAL J. Gen. Virol. 76 (Pt 4), 799-811 (1995)
 MEDLINE 97201539
 PUBMED 9049325
 FEATURES
 source
 5'UTR
 stem_loop
 gene
 CDS
 gene
 CDS
 gene
 CDS
 /location=/Qualifiers
 1..2478
 /organism="Helicoverpa armigera stunt virus"
 /isolate="black mountain"
 /db_xref="taxon:37206"
 /clone="hasvrna2"
 1..282
 1..50
 /note="stem loop which has hexamer sequence, gguaaa, in same place on a stem loop structure on RNA1; putative"
 /function="unknown"
 283..756
 /gene="p17"
 283..756
 /gene="p17"
 /function="unknown"
 /note="has pest characteristics, makes tube-like structures when expressed in bacteria, unstable when expressed in baculovirus, start codon in poor context so is poorly expressed if at all; putative"
 /codon_start=1
 /protein_id="AAC37884.1"
 /db_xref="GI:609504"
 /translation="MSEHTIANSITTPPYTIALIPPEDEAGEMLEMRHSULTTYAE
 PTFGAPTPSPSWETINGVGEKFLDTITSPILHKTSKALTPPLSPSTSLAC
 PNGSIGPRRSSTIPSAGSTSTLQRVLOSIRASASTRSLTASSSSPSSTOR"
 366..2309
 /gene="p71"
 366..2309
 /gene="p71"
 /standard_name="Hasy coat protein"
 /note="coat protein for HasyV, forms T=4 capsids with 240 copies of protein per capsid, upon assembly into particle is cleaved at N/F at residue positions 575-76"
 /codon_start=1
 /evidence=experimental

QY 521 CysSerIleValThrLysThrTyrgInglYTrpGluValThrAsnValAsnThrPro 540
|||||
Db 1926 TGCGATTCGTCTACTAAGACCTACAGGGTTGGGAAGCGCTCAGCAACGTCAACACGCTT 1985
QY 541 PheGlyInPheAlaHisAlaGlyLeuLeuLysAsnGluIleLeuCysLeuAlaAsp 560
|||||
Db 1986 TTGGCCAAATTCGGGCAACGGGGCTCTCAAGAAATGAGGAGATCGTCTGCTGGCGGAC 2045
QY 561 AspLeuAlaThrArgLeuThrGlyValTyProAlaThrAspAsnPhaAlaAlaVal 580
|||||
Db 2046 GACTGGCCACCCCTCTCAAGAGGTGTACCCGCCACATGCACTTCGGCGCGCTT 2105
QY 581 SerAlaPheAlaAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIle 600
|||||
Db 2106 TCTGCTTCGCCCGCAACATGCTCTCTCGTCTGAAGTGGAGGCAACCTCTCCATC 2165
QY 601 IleLysSerValGlyIuThrAlaValAlaGlyInSerGlyLeuAlaLysLeuPro 620
|||||
Db 2166 ATCAAGTCCGTGGCGAGACTGCCGTGGCGGCTCAGTCCGGCTCGCAACCTACCC 2225
QY 621 GlyLeuLeuMetSerValProGlyLysIleAlaAlaArgValAlaArgAlaArg 640
|||||
Db 2226 GGAATGCTAAATGAGTGTACACAGGAGAAATGCGCGGTGTCCGCGCGCGCGAGCGCG 2285
QY 641 ArgArgAlaAlaArgAlaAsn 647
|||||
Db 2286 CGCGCGCGCGCTCGTCCCAAT 2306

RESULT 7
ARI25098
LOCUS ARI25098 2479 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 51 from patent US 6177075.
ACCESSION ARI25098
VERSION ARI25098.1 GI:14111160
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2479)
AUTHORS Christian,P.Daniel., Gordon,K.Henrich,Julius, and
Hanzlik,T.Nelson.
TITLE Insect viruses and their uses in protecting plants
JOURNAL Patent: US 6177075-A 51 23-JAN-2001;
FEATURES
source location/Qualifiers
1..2479
BASE COUNT 516 a 855 c 607 g 501 t
ORIGIN
Alignment Scores:
Pred. No.: 2.23e-210 Length: 2479
Score: 3361.00 Matches: 647
Percent Similarity: 99.85% Conservative: 0
Best Local Similarity: 99.85% Mismatches: 0
Query Match: 99.61% Indels: 1
DB: 6 Gaps: 0

US-09-677-653a-50 (1-647) x ARI25098 (1-2479)
QY 1 MetGlyAspAlaGlyValAlaSerGlnArgProHisAsnArgGlyThrArgAsnVal 20
|||||
Db 366 ATGGGAATGCTGAGAGGCGCTCACAGCACTACAAACCTCGCGGAGACCCGTAACGCTT 425
QY 21 ArgValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArgArgArgThrGly 40
|||||
Db 426 CGGGTCAGCGCCAAACACCGTCACCGTCAATGTAGAGAAACCAACGGCGTCGAGACCGGA 485
QY 41 ArgGlnValSerProAspAsnPheThrAlaAlaGlnAspLeuAlaGlnSerLeu 60
|||||
Db 486 AGGGAAGTTTCTCCCTGCAAAATTTCAACGCTCTGCACAAAGACCTCGCAAGACCTT 545
QY 61 AspAlaAsnThrValThrPhe-ProAlaAsnIleSerSerMetProGluPheArgAsnTr 80
|||||

Db 546 GAGCCAAACACCGTCACTTCCCGGTAACATCTTAGATGCCGGAATTCGGAAATG 605
QY 80 PalaGlyLysIleAspLeuAspSerSerIleGlyTrpTyPheLysTyLeuAs 100
|||||
Db 606 GGCCAAAGGAAGATTCAGACCTCGACTCGAATTCATCGGGCTGATCAATGACCTTGA 665
QY 100 PProAlaGlyAlaThrGlnSerAlaArgAlaValGlyGlnTySerLysIleProAspG 120
|||||
Db 666 CCCAGCGGTGCTACAGATGTGGCGCGCCGTGGGAGAGTCCGAAGATCCCTGACGG 725
QY 120 YLeuValLysPheSerValAspAlaGluIleArgGluIleTyAsnGluGluCysProVa 140
|||||
Db 726 CTTGCTCAAGTTCCTCGTCCAGCAGAGATAGAGATCATATACAGAGAGTCCCGCT 785
QY 140 ValThrAspValSerValProLeuAspGlyArgGlnTrpSerLeuSerIlePheSerPh 160
|||||
Db 786 CGTCACTAGAGTGTCCGCCCCCTCGAGCGCGCCAGTGGAGCTTCGATTTCTCCTT 845
QY 160 ePromeThrArgThrAlaTyValAlaValAlaAsnValGluAsnLysGluMetSerLe 180
|||||
Db 846 TCCGATGTGGAACCCGCTACGTCGCCGTAGCGAACGTCGAGAACAGAGATGTCGCT 905
QY 180 uAspValValAsnAspLeuIleGluTrpLeuAsnAsnLeuAlaAspTrpArgTyValVa 200
|||||
Db 906 CGAGCTGTCAACGACCTCATCGAGTGGCTCAACAATCTCGCGACTGGCTTATGTCGT 965
QY 200 lAspSerGluGlnTrpIleAsnPheThrAsnAspThrThrTyTyValArgIleArgVa 220
|||||
Db 966 TGACTGTGAACAGTGGATTAATCTTACCAATAGACACACAGTACTCGCACCTCCGCGT 1025
QY 220 lLeuArgProThrTyTyAspValProAspProThrGluGlyLeuValArgThrValSerAs 240
|||||
Db 1026 TCTACGTCCAACTCAGACGTTCAGACCCGACAGAGGGCCCTTGTGCAAGTCTGAGA 1085
QY 240 pTyArgLeuThrTyTyAsaIleThrCysGluAlaAsnMetProThrLeuValAspG 260
|||||
Db 1086 CTACCGCTCACTTATTAAGGAGATGATGGAAGCCAAATGCCAACACTCGTCGACCA 1145
QY 260 nGlyPheTrpIleGlyGlnTyTyAlaLeuThrProThrSerLeuProGlnTyAspVa 280
|||||
Db 1146 AGGCTTTGGATCGCGCGCCAGTACGCTTCACCCGCACTACGCGAGTACGACGT 1205
QY 280 lSerGluAlaTyTrAlaLeuHisThrLeuThrPhaAlaArgProSerSerAlaAlaAla 300
|||||
Db 1206 CAGGAGGCTACCGCTGTGACACTTGTGACCTTGCACAGACCAATCCAGCGCTGCAC 1265
QY 300 uAlaPheValTrpAlaGlyLeuProGlnGlyTyThrAlaProAlaGlyThrProAlaTr 320
|||||
Db 1266 CGCGTTTGTGTGGGACAGTTTGGCCACAGGTGGGACAGTCCGCTCGAGCACTCCAGCCTG 1325
QY 320 pGluGlnAlaSerSerGlyGlyTyTyLeuThrTrpArgHisAsnGlyThrThrPheProAl 340
|||||
Db 1326 GGACAGAGGACCTCGGGGTGGCTACCTGACCTGGCGGCACAAACGTTACTTCCAGC 1385
QY 340 aGlySerValSerTyValLeuProGluGlyPhaAlaLeuGlnArgTyTrAspProAsnAs 360
|||||
Db 1386 TGGCTCGTTAGCTTACGTTCTCCCTGAGGGTTTGCCCTTGAGGCTACAGCCGAACGA 1445
QY 360 pGlySerTrpThrAspPheAlaSerAlaGlyAspThrValThrPheArgGlnValAlaVa 380
|||||
Db 1446 CGGCTCTTGAGCGCACTTCCGTTCCGAGAGACACCGTCACTTCCGGGAGGTGGCGGT 1505
QY 380 lAspGluValValThrAsnAsnProAlaGlyGlyGlySerAlaProThrPheThrVa 400
|||||
Db 1506 CGAGAGAGTGTGTGACAAACAAACCCCGCGGCGGAGGAGCGCCCGCACCTTCACCGT 1565
QY 400 lArgValProProSerAsnAlaTyTrpAsnThrValPhaArgAsnThrLeuGluTh 420
|||||
Db 1566 GAGAGTCCCTTCAACAGCTTACACCAACCCGTGTTTAGAACAGGCTTTAGAGAC 1625
QY 420 rArgProSerSerArgArgLeuGluLeuProMetProProAlaAspPheGlyGlnThrVa 440
|||||
Db 1626 TCGACCTCTCTGTAGGTGCGAATCCCTATGCCACCTGTGCTACTTTGGACAGAGAGGT 1685

QY 440 lalaasnaasprolyslleaglinslerleuleuylsgluthrleucllycystyrleuva 460
 Db 1686 cggcaaacacccgaaagatcgacagctgctttaaagaaacattggctgctatttggt 1745
 QY 460 lhisserlysmelaargasnaarprovalpheglnleuthrprolasersestheglialava 480
 Db 1746 ccactccaaatgcgaaccccgcttttcgactgacgcccagccagctcttggccgct 1805
 QY 480 lserpheasnaasproglytyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 500
 Db 1806 ttctcttcaacatccgggttatgacgacacccgacactccggactacactggactccg 1865
 QY 500 gaaspeaspeaspglnasmseserthrallavalalaahisphearsestheulserhis 520
 Db 1866 tgactatctgcacgaacacatgctcaacgctgagccacttccgctacacttccacac 1925
 QY 520 rcysserllevalthrlyserthrlyserthrlyserthrlyserthrlyserthrly 540
 Db 1926 ctgcactatctgctactaagacactaagacaggttggaagggctcagcaacgctcaac 1985
 QY 540 ophegllyglnphealhisalaglyleuleuylsasnlglnlleucysleuualas 560
 Db 1986 ttctggcccaattcgccgacgcccgcctctcctcaagaatgagagactcttgcctcc 2045
 QY 560 piaspleualatthargleuthrlytyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 580
 Db 2046 cgacactgcccacccgctctcagagctgctaccccgccactgacacactgcccgcct 2105
 QY 580 lseralalpheal 600
 Db 2106 ttctgctcttgcgcccgaacactgctgctcctccgctggaactcgagagcaacgctcc 2165
 QY 600 eilelysservalgllythrallavalaglialalaglnserglyleualalysleupr 620
 Db 2166 catcaagctcgttggcgagactgctcgccggcgctcactccggcctccggaagctacc 2225
 QY 620 oclyleuleumetservalproglysllealalalalalalalalalalalalalal 640
 Db 2226 cggactgctaatgactgacacaggaagattggcgccgctgctcggcgccgagacggc 2285
 QY 640 gargargal 660
 Db 2286 ccgcccgcgcccgtgctgctgctgctgctgctgctgctgctgctgctgctgct 2307
 RESULT 8
 LOCUS S43937 2448 bp RNA linear VRL 08-MAY-1993
 DEFINITION capsid protein [Nudaurelia capensis omega virus, Genomic RNA, 2448 nt].
 ACCESSION S43937
 VERSION S43937.1 GI:255203
 KEYWORDS
 SOURCE Nudaurelia capensis omega virus.
 ORGANISM Nudaurelia capensis omega virus
 VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tetraviridae; Omegatetravirus.
 REFERENCE 1 (bases 1 to 2448)
 AUTHORS Agrawal, D.K. and Johnson, J.E.
 TITLE Sequence and analysis of the capsid protein of Nudaurelia capensis
 JOURNAL Virology 190 (2), 806-814 (1992)
 MEDLINE 92391097
 PUBMED 1519360
 REMARK Genbank staff at the National Library of Medicine created this
 entry [NCBI glibseq 112777] from the original journal article.
 Map location: segment RNA2.
 FEATURES
 source 1, 2448
 location/Qualifiers
 /organism="Nudaurelia capensis omega virus"
 /db_xref="taxon:12541"
 gene 367..2301

CDS
 /gene="capsid protein, CP"
 367..2301
 /gene="capsid protein, CP"
 /note="this sequence comes from Fig. 5; Protein sequence from Fig. 3 does not match conceptual translation; CP"
 /codon_start=1
 /product="capsid protein"
 /protein_id="AA023198.1"
 /db_xref="GI:255204"
 /translation="MDSNSASGKRSSRNRYRIANTVNVAPKORAGRARRANNID
 NVYAAOELGOSLDANVTPEFNVATMPERFARMGAKLIDDSIGWYFKYDIPAGAT
 ESARAVGERSKIPDGLVRSVVAEIREITNECPYSDASTPLDQNSLSTISPMF
 RTAFRAVAVDNKEISLDVTNDLIVNLNLASRVDSGQFITSDDTFVRLRLV
 HPTYDLPDTEGLRCSQDYLRTKRSITCEANMPTLVQGEWIGHYALPTATQNA
 VEGSGEHEFNVTRPGIAGVTLRTWASMPGASGDPAMIPDSITQOMRHGFDA
 PGVITVTYTPRGVIMQYFDTTENNNGNFPDQVYFGQIGAGATNATITTAFTV
 LTIATTTTSAANVINFRNLDAETTTAASNSEVPLPLPGGOTAPNNPKTEOTLVKDTL
 GSYLVHSKRNVFQOLITPSSGALSTFPRGDRULDPFRGFGITSDSDVNNSTAVCH
 FSLSSCSITVKTTOGMEGVNTVPTPOGFHSGLKNDITCLADLADLATVLTGVG
 AIDNFAAVALAFNAAMLTSLVLSSEATTSVYIKELGNQAGLANOGLARLPGLLASIPGK
 IAAVRARDRRRRAAMNNN"
 BASE COUNT 547 a 715 c 563 g 623 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,97e-135 Length: 2448
 Score: 2207.00 Matches: 428
 Percent Similarity: 76.07 Conservative: 71
 Best Local Similarity: 65.24 Mismatches: 139
 Query Match: 65.418 Indels: 18
 DB: 14 Gaps: 6
 US-09-677-653a-50 (1-647) x S43937 (1-2448)
 QY 1 MetGlyAspAlaGlyValAlaSerGlnArgProHisAsnArgGlyThrArgAsnVal 20
 Db 352 CTGGGAACCTTTTGAATGGACAGTAACAGCTCCGCGGACGAGATCGCCAAATG 411
 QY 21 ArgValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArg-----ArgArg 38
 Db 412 CGGATTGCAACCAACCGGACGTCGCTCCGACACAGCCACAGCAGCGGCGCCCA 471
 QY 39 ThrGlyArgGlnValSerProProAspAsnPhierThrAlaAlaAlaGlnAspLeuAlaGln 58
 Db 472 GCCCGTTACAGTGCACCAACAGATGTCACAGTGCAGCAGTGCAGCAATCGGCCAA 531
 QY 59 SerLeuAspAlaAsnThrValThrPheProAlaAsnIleSerSerMetProGluPheArg 78
 Db 532 TCCTTGTACGCAACAGCTTATACCTTCCCAACCAATGTTGCAACCAATGCGATTCG 591
 QY 79 AsnTrpAlaGlyGlyLysIleAspLeuAspSerAspSerIleGlyTrpTyrrPheLysTyr 98
 Db 592 AGCTGGCGCGGTGTAACATGTCATCGACGACAGACTCCATTGGTGTGATTTCAATAC 651
 QY 99 LeuAspProAlaGlyAlaThrGlnSerAlaArgAlaValGlyLysIleSerLysIlePro 118
 Db 652 CTGACCGCTGCTGGCGCCACGAGTCCGCTCGGCGGTGAGTACGCAAAATATACCT 711
 QY 119 AspGlyLeuValLysPheSerValAspAlaGlnIleArgGlnIleLysArgGlnGlyLys 138
 Db 712 GAGCGGCTGCTCAAAATTTCTGTGACGCGGCAATTAAGAGATTAACATAGGAGTGC 771
 QY 139 ProValValThrAspValSerValProLeuAspGlyArgGlnTrpSerLeuSerIlePhe 158
 Db 772 CTTACCGCTTCGACCTTCTATACCATGTCGACGCGCCCAATGAGCTTGTCTATATC 831
 QY 159 SerPheProMetPheArgThrAlaTyrValAlaValAlaAsnValGlyLysArgGlnGly 178
 Db 832 TCTTACCGTATGTTTGCACCTGCTTACTTGGTGTGCGACAGTGTATATAAGAGATA 891
 QY 179 SerLeuAspValValAsnAspLeuIleGlyLysPheAsnAsnLeuAlaAspTrpArgTyr 198
 Db 892 TCCTTTGACGTACGAAACGATCTTATCGTTTGTGCTTAACATCTCGTAGTTGGCGGTGAC 951

Oy	552	AsnGluGluIleLeuCysLeuAlaAspSerPheuAlaThrArgLeuThylValTyrPro	571
Db	2005	AATGACGAGATTCTTGGCTTGCACACACACTTACCTACCGCTTACAGCGCTTATGCT	2064
Oy	572	AlaThrAspAsnPhaAlaAlaAlaValSerAlaPheAlaAlaAsnMetLeuSerVal	591
Db	2065	GCGAGCGATTAATTTGCTGTGCGGGTCTCCGCTTCCGCCAAATATGCTCATCCCTGC	2124
Oy	592	LeuysSerGluAlaThrSerSerIleIleIysSerValGlyIuThrAlaValGlyAla	611
Db	2125	TTGAAGTGTGGAAGCTACCACTTCAGTTATTAAGAGCTCGCAATCAAGCTACTGGACTT	2184
Oy	612	AlaGlnSerGlyLeuAlaAlaLysLeuProGlyLeuLeuMetSerValProGlyLysIleAla	631
Db	2185	GCCAAACGAGGGCGCTTGCCTTACCGGGTCTTCTGTGCTGATATCGCGGCAAAATGCT	2244
Oy	632	AlaArgValAlaArgAlaArgArgAlaArgArgAlaAlaArgAlaAsn	647
Db	2245	GCACCTGTCCGCGCTAGACGTGACCGCTGCAGCTGCTGCACGAAATGAAAC	2292
RESULT 9			
LOCUS	AF062037	2482 bp	mRNA linear VRL-23-JUL-1999
DEFINITION	Thosea asigna virus capsid protein precursor (Tav-Cp) mRNA,		
ACCESSION	AF062037		
VERSION	AF062037		
KEYWORDS	AF062037.1 GI:3861489		
SOURCE	Thosea asigna virus.		
ORGANISM	Thosea asigna virus.		
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage: Tetraviridae; unclassified Tetraviridae.		
AUTHORS	1 (bases 1 to 2482) Pringle,F.M., Gordon,K.H., Hanzlik,T.N., Kalmakoff,J., Scotti,P.D. and Ward,V.K.		
TITLE	A novel capsid expression strategy for Thosea asigna virus (Tetreviridae)		
JOURNAL	J. Gen. Virol. 80 (Pt 7), 1855-1863 (1999)		
MEDLINE	99350019		
PUBMED	10423156		
REFERENCE	2 (bases 1 to 2482) Pringle,F.M., Kalmakoff,J. and Ward,V.K.		
AUTHORS	Direct Submission		
TITLE	Submitted (29-APR-1998) Microbiology, University of Otago, 700 Cumberland Street, PO Box 56, Dunedin 9001, New Zealand		
JOURNAL	3 (bases 1 to 2482) Pringle,F.M., Kalmakoff,J. and Ward,V.K.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (12-NOV-1998) Microbiology, University of Otago, 700 Cumberland Street, PO Box 56, Dunedin 9001, New Zealand		
TITLE	Sequence update by submitter		
REMARK	On Nov 13, 1998 this sequence version replaced gi:3859469.		
COMMENT	Location/Qualifiers		
FEATURES	1..2482		
source	/organism="Thosea asigna virus"		
	/specific_host="Setothosea asigna"		
	/db_xref="taxon:83810"		
	/note="old name for host is Thosea asigna"		
gene	1..2482		
	/gene="Tav-Cp"		
CDS	88..2361		
	/gene="Tav-Cp"		
	/note="forms T=4 capsids with 240 copies of the large and small capsid proteins per capsid"		
	/codon_start=1		
	/evidence="not_experimental"		
	/product="capsid protein precursor"		
	/protein_id="AAC97195.1"		
	/db_xref="GI:3861490"		
	/translation="MIVLKRLTSMNRVLEONLKNLKIKPIGFMRSNGSGROIVAA LANLIKPCVNSSLRTEHPSRSWKRGTIOEYQALMEIKELIEIVPOLQEDTLISF ENPEEGSSNLKDKSMRPRPSRRRLRGAPOLGVRAIGRSGLTLCGVENPEPTIVA GLVQEAEGGLNKATKAPGKTRTRRRRRNGKRIQNPQPSISGAAPMSKIPALRGTL		

Db 1980 ATGTCAATGACCTCTTTGGCCCTTGACGCTCCGTTATTAAGCGGAGATGCGCTT 2039

QY 531 pgluglyvalthrAsnValasnThrProPheglyglnPheAlaHisAlaGlyLeuLeu 551

Db 2040 TGAAGGCGCTCCGAGAGAAATCCCGGAGACCTTTGGCCACCTCCCTCCAA 2099

QY 551 saanglgluileucysleuAlaaspAspIleuAlaThrArgLeuThrGlyValTyrPr 571

Db 2100 GGACGAGGTTGCTTACCTATTCACAGACTGTGATGAGACCTTGACCATTTGCTATGCC 2159

QY 571 calathraspAsnPheAlaAlaAlaValSerAlaPheAlaAlaMetLeuSerSerVa 591

Db 2160 CCGGCTTATATAGCTGGGGGTTGATGTTCTCGAAAGCTTCTATGTCGGAAGAT 2219

QY 591 lleuYsSerGluAlaThrSerSerIleIleYsSerValGlyGluThrAlaValGlyAl 611

Db 2220 ACCCAAGCCCTCAGAACCTCCAGC-----AATGTTGACAGACTGTTCCAGGCTG 2270

QY 611 a-----lAlaGlnSerGlyLeuAlaValSerProGlyLeuLeuMetSerValProGl 628

Db 2271 TGTAGAGAGAGCTCAGCGCCGCTCAG----- 2298

QY 628 ylyslleAlaAlaArgValArgAlaArgAlaArgArg 642

Db 2299 -AAAGTGTATGAAGACTGTCGTGGCGCGAGAGACTATGAGA 2340

RESULT 10

AF461742 5698 bp ss-RNA linear VRL 20-FEB-2002

DEFINITION Euprosteria elaeasa virus, complete genome.

ACCESSION AF461742

VERSION AF461742.1 GI:18766345

KEYWORDS

SOURCE Euprosteria elaeasa virus.

ORGANISM Euprosteria elaeasa virus.

REFERENCE 1 (bases 1 to 5698)

AUTHORS Zeddam,J.-L.A., Pringle,F.M., Gordon,K.H., Ward,V.K., Luke,B.T., Goralenya,A.E. and Hanzlik,T.N.

TITLE Genome organization of Euprosteria elaeasa virus defines it as a member of a new group of insect RNA viruses

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 5698)

AUTHORS Zeddam,J.-L.A., Gordon,K.H. and Hanzlik,T.N.

TITLE Direct Submission

JOURNAL Submitted (20-DEC-2001) Insect Virology, IRD, Route de Montpellier, Saint-Christol-lès-Ales 30380, France

FEATURES

source

1..5698

/organism="Euprosteria elaeasa virus"

/db_xref="taxon:186165"

/note="Isolated from larvae of Euprosteria elaeasa, Dytar (Lepidoptera: Limacodidae)- a major defoliating pest of oil-palms in several Latin American countries"

95..3868

/codon_start=1

/product="putative replicase"

/protein_id="AAI78970.1"

/db_xref="GI:18766346"

/translation="MEASNPVIAPRSLSEAMLEERASVRODLOSRLVRLSGAPRTL APOAPEKGRDASAQAASIVELRDVAEAGLEHKSIDPELOGEYVCCIIHGDH PIHPSTGNKKKSVQROFTTSMSVDAIOVALREDPVLYRDIALKTTTIDGSLAG FWERLLOTRPCQVEKVSRLRLRELPQKLVDSSELSDLVEIKTNATASGP PYWRSKRALPMDLCVPLRYOHIVDGLATLKNKHPLEFLAEKKNKDRLEVEKLG EKTREYFSPHFLSALVSYSQSFKALTIMTEESTSYNAYSGEKNKGAEGLIVAR GAGEVGEHPRIACGDDTDIYRQKGRILYICPDKODGSDVDETTIKAAVDYVVA HRIOYPLAKGFWLEKLMWEMATSPEDICTGTLVYRKORNGSLMGVGTLLPDIYK SALAYKDMADOLLEBDLELAKKKAIFKEKHGLYIKETGNPAVTDVDPIDELMT FOKFLGLOIKVYKADVDKVFVPLTSPDWLSMWVTPARGLHKESTYMRQRTLEDR ARGLLVYTGAAFKQARLELMSVINEPSEIYVCMKVOEGGGRGAPSYALTELDGNEEF PVSDFGPDIAATFRLYRSDDTGMDMPFEDAVDTLIAAYKRLMDRIRIYKVEGNSYLM EYAAVETGIMNEMPAVAVLPPKPSIOKPREFTVAENSTGSLVYKVKVDPSA GTLETFIQKFMESVVKQIEELRAGQEKERTITIMGAALPAEAIKKMGVNDLKRTEK

CDS

AALEPHEAGDLYLSTNPVLGHDPDSERGEARAEVHKVILGKHSPTERRAKREKPEVHD TGEFVKITTPAGGVEPTQCEPOTPWYSLCRSVSAQVYVTRTERPGDHNRPVRA VVLOKMRMLKPPQWDEWLVTGSLARAEELASDVALELVQERKQOPDATTSAED ABIGDTIQLDGRHVEFLIGERFERNVIGTGHOLTDROGLMWOPASGYVVAAP PKRTSGARSRTLASVKVSVRANADLRATGAGEPFRDPVLSNLKKIKNKEARL HSEEDLSREPEYORGPSSEFEDSRADDESDAORHVRVVOGLSTKIDNKEARL ADRARDREPSLIDKEAGVPOVSGSRPYSRSLRPNPVOGLEOVASLOAOITTTI ACVRAPRPGCARKSQNMRRRGESAPORCGAGSGGASQNKPRPKSM"

CDS

3336..3606

/codon_start=1

/product="capsid protein precursor"

/protein_id="AAI78971.1"

/db_xref="GI:18766347"

/translation="MKIRRLSLMNQKILEANRLNKEGHPLEAFMRTACGRKSIAMLA NMIVCPTCNLSLNKREKRTDWRGTLDELEUMLWKRELRLCPCPDHLLIDQAGI PPPESSNSKSVHFHRNSRLPESQOLPOGAGRLVYCGEENPGPSVAYLRQLE VVAGLAKTGSPEARNGKARRRRRRYKCGNQNQPAVYAGALATVPALRGLA MSEIPSSVGMLETNLDPCGEYKSTLDYGVKYPDGAIPMSVCGREFETTRHGPAN MYAPLDGTMWTIGIIRTNLMRTPPLILIAMRNAEISDSLDEVARSPNNVGEIELAS HPNMYETLOEGLYMSIQRTALICVPPSDPVSPIYDPRMTSGFTYSSNTPSLIN QGAVVVAQFNPENTENKEIVPHSESGETPMGSRVSLSTQSIVTITIQGVGAQVTE TVIANRGPAINSSVLPVAVTQGYRELNGDTIVAGSLRYQTERDSGGTYVLYORST GTTFNIGPSTTTAAPSCTSTASGVFLIEDVLAGRVNMLTLPFTOSDLMQOTPK TCVPQIKREGGFYRODVPQVYNTMPASRYAVRFVSGAITDLSLNSVGTIRDTADS NYGFALCHMTSILPACAPRIKAKEMREAVRGNRSPMGLFATSTPPKDEVALTISRYM DLPEPFAMPLAYNGWGLMFSKIINIYAKVKVLTASNVADVANCVEDQAQATRYVE SVGRRGANRDMTLOO"

3336..3785

/product="small putative peptide"

/note="not encapsitated in the mature virion"

mat_peptide 3786..5420

mat_peptide 3786..5420

mat_peptide 5421..5603

BASE COUNT 1385 a 1357 c 1616 g 1340 t

ORIGIN

Alignment Scores:

Pred. No.: 5,02e-17 Length: 5698

Score: 407.50 Matches: 157

Percent Similarity: 39.69% Conservative: 97

Best Local Similarity: 24.53% Mismatches: 287

Query Match: 12,08% Indels: 99

DB: 14 Gaps: 22

US-09-677-653a-50 (1-647) x AF461742 (1-5698)

QY 32 ArgArgasnglnArgArgArgThrglyArgGlnValSerProProaspAsnPheThrAla 51

Db 3876 AGGAGCGCGCGCTGCGCTTACAGCCGTTAAAGCAGCGGAATCCCAAT----- 3923

QY 52 AlaAlaGlnAspLeuAlaGlnSerLeuAspAlaAsnThrValThrPheProAlaAsnIle 71

Db 3924 -----CAGCCTTTTGG-----GTACCGCTGGCGCGCTCTA 3956

QY 72 SerSerMetProGluPheArgAsnTrpAlaGlyGlyLysIleAspLeuAspSerSer 91

Db 3957 ACGGGGGTTCCCGCGCTTGGACACTCTCGCATGTGCGAACCCATCTCACCTTGTGTCG 4016

QY 92 lIlgLItyrpyrThelystyrIleuaspproAlaGlyAlaThrGlnSerAlaArgAlaVal 111

Db 4017 GTAGGTTGGTGAAGACCAACTTGACCATGTGAGTATAGTCGACCTA----- 4070

QY 112 GlyLtyrSerIleProaspGlyLeuValIlyspHeserValaspAlaGluIleArg 131

Db 4071 ---GACTATGTAAGGCGCTGATGTGGCATCCCATGCGTGTGTGTCAGTTAAG 4127

QY 132 GluIleTyraGlnGluGlyProValIvalThrAspValSerValProleuaspglyArg 151

Db 4128 GAGACTTTTCAGATCCGACATCTCTGGGGCCCAACCCATGATGTCCTTTGATGACACC 4187

QY 152 GlnTrpSerIleuSerIlePheSerPheProMetPheArgThrAlaIyValAlaValAla 171

Db 4188 ATGTGACCCCTTGATTAATCCGACTAATTTATGAGGACTCCGTTCTTGATTCGA 4247
 Oy 172 AasnValGluAsnLysGluMetSerLeuAsnProValValAsnAspLeuLeuIuProLeuAsn 191
 Db 4248 GACATGAGGAAATGCTGAATATCCGATTATCTCTAGATGAGTACTGCGCTTTCAAC 4307
 Oy 192 AasnLeuAlaAspTrpArgTrpValValAspSerGluGlnTrpLeuAsnProAsp 211
 Db 4308 AATTTATGTCGGGAGATTGAATTAGGCTGCACCCGAAATGGGTGAGACCTTGACGAG 4367
 Oy 212 ThrTrpTrpValValArgValLeuArgProThrTrpAspValProAspProThr 231
 Db 4368 GGCCTGTAATGG---TCCATCCAGCGCTGAGCGCTTTAATAGCGCTCCACCTCCCACT 4424
 Oy 232 Glu---GlyLeuValArgThrValSerAspTrpArgLeuThrTrpLysAlaLeuThrCys 250
 Db 4425 GATACGGGGGTTTCCCTATATCACTGATTTCCGGATCACAGGTTCCGGATTTTACTGTG 4484
 Oy 251 GluAlaAsnMetProThrLeuValAspGlnGlyPheTrpLeuGlyGlyGlnTrpAlaLeu 270
 Db 4485 TCTTCCAAACACCCCACTCTTATTAATCAAGGATGGTGGTGCCACAGTTT----- 4538
 Oy 271 ThrProThrSer-----LeuProGlnTrpAspValSerLys----- 282
 Db 4539 AACCCGAACACTGAGAACAGAGATAGTACCCATTTCTGAGCTGGGAGACTCCGATG 4598
 Oy 283 -----AlaTrpAlaLeuAsnThr-----LeuThrPheAlaArgPro 294
 Db 4599 GGGTGGTCTGGGTTTATTCCTATCACTACAGTACAGATAGCTTATATTCACCA 4658
 Oy 295 SerSerAlaAlaAla-----LeuAlaPheValTrpAlaGlyLeuPro 308
 Db 4659 GGGGTGGAGGTGCCAGAAATCTTCACAGTACATGCCAAATAGAGGGCTCTGTAAAT 4718
 Oy 309 GlnGlyGlyThrAlaProAlaGlyThrProAlaTrpGlnAlaSerSerGlyGlyTrp 328
 Db 4719 AATTAAGCTGGTTCTCCCTGCTCACTCAAGGTTTACCGT-----TTT 4760
 Oy 339 LeuThrTrpArgHisAsnGlyThrThrPheProAlaGly---SerValSerTrpValLeu 347
 Db 4761 TTA-----AATGGACACAGTATGCAAGTTGGGAGACTCCCTCCGATAT----- 4802
 Oy 348 ProGlnGlyPheAlaLeuGlnArgTrpAspProAsnAspGlySerTrpAspPheAla 367
 Db 4803 -----CAGTTTACGGAGCGAGAC 4820
 Oy 368 SerAlaGlyAspTrpValThrPheArgGlnValAlaValAspGluValValValThrAsn 387
 Db 4821 TCAGCGCGGTATACAGTC---TATCTCGAAAGCTCACTGATGTGATGACACTTTTCAAAAT 4877
 Oy 388 AsnProAlaGlyGlySerAlaProThrPheThrValArgValProProSerAsnAla 407
 Db 4878 -----ATTGGCCCGCGCTCGACGACAAACGCGCCCATCTGGCACT 4919
 Oy 408 TyrTrpAsnThr-----ValPheArgAsnThrLeuLeuGlnThrArgProSer 423
 Db 4920 TCACAGCGCTTCTGGGCTTTCATCTAGAAAGTGAATGTTTGGCCGACGCTGTGAAT 4979
 Oy 424 SerArgArgLeuGlnLeuProMetProProAlaAspPheGlyGlnThrValAlaAsnAsn 443
 Db 4980 ATG-----TTAACTTACCAGCCCTTTTACCAGCTGTGATCTGTGCAACAGACC 5027
 Oy 444 ProLysIleGlnGlnSerLeuLeuLysGlnThrLeuGlyCysTrpLeuValHisSerLys 463
 Db 5028 CCAAGAGCTGTGTGTTCCAGATCAAGAG-----GGTGGTTTATGTGGCCAGGAT 5081
 Oy 464 MetArgAsnProValPheGlnLeuThrProAlaSerSerPheGlyAlaValAspPheAsn 483
 Db 5082 GTTTGGCAACCGCTATATATGACCCAGCCAGCCGTTATGCCCGCTTGATTTGG 5141
 Oy 484 AsnProGlyTrpGlnArgThrArgAspLeuProAspTrpTrpGly---IleArgAspSer 502
 Db 5142 TCG---GGAGCAATTAATTAGATAGCCTGAATTTCTTGTGAGAACTATCCGCGATACA 5198

Oy 503 PheAspGlnAsnMetSerThrAlaValAlaHisPheArgSerLeuSerHisSerCysSer 522
 Db 5199 GCAGACTCTAAGTATGAGGTTTGGCGCTGTGCATGACCTCCATACCTGGCGTCGCC 5258
 Oy 523 IleValThrLysThrTrpGlnGlyTrpGlnGlyValThrAsnValAsnThrProPheGly 542
 Db 5259 CCTTTATTAAGCTGAGATGCGCTTGAAGCGGTGCCAGTCCGAAACTCTCCGTGGGG 5318
 Oy 543 GlnPheAlaHisAlaGlyLeuLeuLysAsnGlnGlnIleLeuCysLeuAlaAspLeu 562
 Db 5319 GTTTTGGCCACTCCACCTCCCGCCCAAGGATGAAGTTGCTGACCATATACAGCAGGTC 5378
 Oy 563 AlaThrArgLeuThrGlyValTrpProAlaThrAspAsnPheAlaAlaValSerAla 582
 Db 5379 ATGCACTTGAACCTTTTCCATGCCCTTCCCTATTAAGGTTGGGCTGTGATTTCTCA 5438
 Oy 583 PheAlaAlaAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIleLeuLys 602
 Db 5439 AAAATCATCATATAGTACGCAAGGTTCCGAAAGTCTCCGACAGCAGCAACGTTGCT 5498
 Oy 603 SerValGlyGlnThrAlaValGlyAlaAlaGlnSerGlyLeuAlaLysLeuProGlyLeu 622
 Db 5499 GATGTTGTTGCTAACTGCTGGAGAGAGCCCAAGCTGGTGGCACC----- 5543
 Oy 623 LeuMetSerValProGlyLysIleAlaAlaArgValArgAlaArgAlaArgArg 642
 Db 5544 -----CGAGTTATAGAGATGTGGGGGGCGCCGAGGGCTATGCCA 5585

RESULT 11
 LOCUS AF102884 6625 bp RNA linear VRL 30-MAY-2001
 DEFINITION Nudaurelia capensis beta virus complete genome.
 ACCESSION AF102884
 VERSION AF102884.1 GI:4028586
 KEYWORDS
 SOURCE Nudaurelia capensis beta virus.
 ORGANISM Nudaurelia capensis beta virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Tetraviridae; Betateiravirus.
 REFERENCE 1 (bases 1 to 6625)
 AUTHORS Gordon, K.H., Williams, M.R., Hendry, D.A. and Hanzlik, T.N.
 TITLE Sequence of the genomic RNA of nudaurelia beta virus (Tetreviridae) defines a novel virus genome organization
 JOURNAL Virology 258 (1), 42-53 (1999)
 MEDLINE 99263183
 PUBMED 10329566
 REFERENCE 2 (bases 1 to 6625)
 AUTHORS Gordon, K.H.J., Williams, M.R., Hendry, D.A. and Hanzlik, T.N.
 TITLE Direct Submission
 JOURNAL Submitted (30-OCT-1998) CSIRO Entomology, GPO Box 1700, Canberra, ACT 2611, Australia
 FEATURES
 source Location/Qualifiers
 1. 6625
 /organism="Nudaurelia capensis beta virus"
 /db_xref="taxon:85652"
 93..5870
 /note="replicase"
 /codon_start=1
 /product="RNA-dependent RNA polymerase"
 /protein_id="AAC97509.1"
 /db_xref="GI:4028587"
 /translation="MEDAKSOLRVLDNOERAKAFOLEFIASVETLEDADKEYGMMF
 RSGTRPSTHKLALDLRYAERKDLRRHKVKNPVYLEIGSVESRYAVAVOTRDERVIG
 CTESPDARDNLNHNKIGYEATRTKIDGDAALLAGIPDTFCYVGFSGECOSPLATA
 CHSLYPPDESWSIMDVAKGMLRGTHYIYAMHLPVELLTITDNDNIEGISTIEET
 GALPCTKRRAKIFSGYNDGSAVYVADHNMWGLKHKGVDPYGFSLIDIQDQKGMH
 TKLRTGHSNGSGITVFPPLSLGLIWPVNIYKIMPKAKIEPYIYDIDKKYGVGV
 YVGRVOSGSKITLAEIYQITRLRLIRITNGVHEKWTWITADODIERLAKYSIMFK
 NVERAVSEKALIMRAOKCKSEKQALLPYMRYIANNFOOKFOIDERYVRRYIECIK
 AQPWTHADKYVNCETKRNPVAVGVGRKNNHLATTTGELRELORETPSANEDPRAKAH
 HSAHADLDIYAEGLIDSAKRAAGKOSLATITDQAOVLKRTCECCANNIDEIYTWG
 PGGGKSRNAAPRFADLOGGLVLCAPRTFLRLDALDESIVHPSRVCTYHNLHVAKES

CDS

GNRPDVIVIDEAETTPACVGTMMHSPSRIVCLGDPHQIGYIDFSRDKDLKPFSS
 IIAECRTFRFTYKCPQDVLNLPFLKLYPDALISFKQLTSIRLTPARSVTPTTH
 AQTIDODQHPSEPTPAHEPOARTRDVVHAGTLEPEALILEKHIVALTIRHNA
 LYIRESEKELVPSLMTPEPSSTYRCPVDMQWDPVAVERENSSSPSCSHHGA
 ITILDELGKIDTVKGVRESAVALPFAHRRVLDLGGPDPTQPMYOTNLNGTIT
 NIKDNOQLHTLVGRYARKINSSREDADYKRIYARLKEIMPFTAPAEQDSCFA
 DAMOKIAERHGVDDIIEDFWNEGORISYHLGOQVMPPTKIKLOQISAEKCANI
 ALSAMVRIIDOMSTSEKIFANQSGDRPMSTIEFIDLEKREPSIDKEEDTVN
 WYSLIVSWRCGRGCEPHLEVEFKRSKRSRSDVSMDSGVTMLNKTETATVTL
 PASGLMALFVGVDFTLAKGDVFLAGNNLYIDAEELRSGSTLANNNKIEXTAVSF
 IGFIVSOAAVTVADVVALTRTYGRSTKNADDLAKYIATADCKLERSPRTILMTAIN
 CATLYGTSKECINYLMDALDAFGHTKMSDLHDPGVMTPEKMDERYSGDDGCP
 ADRTREKQEPGPPGPOQQAQSTQESKTSRPSRTQOPRGDRTREPREGQOR
 SDTRREGVAKSDGESHSDIRGMDSLRSPGRIODEPGRREMSRRDVSNNRSTIR
 DRGROIPTGEPEFDATAGMRDLASTSDASVYLASVYVHHHOOHSGRSPDRSCYRE
 RLEODGDRLDRSPVKIGASREYVLERPRDSADGRTTPSTGCIYRTRELPSDIERH
 SVLQRTHSRESGSGDRAVPTGQRTPEGPGHSSRHPNGRNVYATARELAHIDDD
 RGPGRVGRSNPATHGVDAGVGAAGVADCEPDTRRRHNNHHHAATRVQDNVYA
 IHSQORGDHRDRGSGATVVRSEFGLTESAHQLONDSTHEHPNDAGNADHSV
 PIORNGELLYABEGVPTRVNRNGDYLMTGAMEDTDCNGRLPPGMMWPGYHROOLR
 DRCHDDRVVYINRLLIGQVPTLRSTGGSEPLRGRCDTSGRGASGSLDPSA
 PRIRETGTORIRGPIRDGQDSDPTSLAISRSECGDGLHRENDRCSELEHJLGE
 AAKSETCWRNRSRSPQSCPHREP

CDS

4039..5877
 /note="coat protein: forms T=4 capsids with 240 copies of
 protein per particle"

/codon_start=1

/product="capsid protein"

/protein_id="AAC9510.1"

/db_xref="GI:4028588"

/translation="MDANYOIRPARNNPSOGNRRNNKRRRRRGLKLPVAVATIT
 SPGMAEPANHANTRYNRGTRRGILROMESPMATSEAWTLDLDDGKTSID
 DKIRPDGAIPQSTCCQFRGTGVARRYGLNSTLPLDGGTWPRLVYALPFRFRHLEIT
 TTSNEVEVNTADLDAFANDMNNRDTWTATYPSMVQVNVFTVMTALDTPPT
 QLGSGLESYRLTSSGVAENAPLVNOGVAIVAFOPDEHOKENDIYATVPTQ
 GGTLOQGGSPVYTLMTIGDVEFGGAIPLEPTYMGMPESGQVLPOTANLTFPG
 NNTITTTTPGSPVTGMQFTASNGDITVDAGALYFAGALDSELDNLINISIK
 IPTNNOMQAMPKTIQIOLNETGTFVMDLPGVYVPRFMTATSGVPRMKPRPTV
 VDHRAIGLODTIDSNFAIGVAMTGMTSTVPTPKVTRPREATIAEBSPMGPRSA
 TTPKDDVALTVARMTDHLFPATPERYNGFALFANVAATVATQIPIPVSAAGVANVA
 TDCIESATESVSNSTSERORRRARVGIARARLVORIGLST"

5644..5649

/note="cleavage site for coat protein after assembly"

6514..6625

/note="tRNA-like structure: valine anticodon; no
 pseudo-knot in the aminoacyl stem"

misc_RNA

BASE COUNT 1696 a 1800 c 1783 g 1346 t

ORIGIN

Alignment Scores:
 Pred. No.: 1,45e-15 Length: 6625
 Score: 386.50 Matches: 174
 Percent Similarity: 38.81% Conservative: 86
 Best Local Similarity: 25.97% Mismatches: 295
 Query Match: 11.46% Indels: 115
 DB: 14 Gaps: 31

US-09-677-653a-50 (1-647) x AF102884 (1-6625)

QY 15 ATGAGlyThrArgAsnValArgValSerAlaIAsnValThrValaAsnGlyArgArgAsn 34
 DB 4060 AGACCCGCGAGAAACCCGAGCCAGGCGCGCAACACACACACACACGCT 4119
 QY 35 GlnArgArgArgThrGlyArgGlnValSerPro-----ProAspAsnPheThrAla 51
 DB 4120 CGACGCGAGGCGGCGGTCTTAAACACTCCCGCTAGTCCACCG-----ATTACACGC 4173
 QY 52 AlAlaIeGlnAspLeuAlaGlnSerIeUAspAlaAsnThrValThrPheProAlaAsnIle 71
 DB 4174 CCGGCGCGAGTGGCCGACACCCGCGACACGCGACACAC-----4212
 QY 72 SerSerMetProGlnPheArgAsnTrpAlaIleGlyGly-----LysIle 85
 DB 5179 TCGGAAGTGAATCTGACGAGCATC-----5202

DB 4213 ---CGGTCACACCGCGCTCGGACACGCGGTAGGCGGTTAAGCAAGCATGATGAGAGT 4269
 QY 86 AspLeuAspSerIleAspSerIleGlyTrpTyPheIleSyrLeuAspProAlaGlyAlaThr 105
 DB 4270 CCATGCGACACGACATCAGAGGATGATTCACGACTATCTAGACCCGCGAGGAATAC 4329
 QY 106 GlnSerIleArgAlaValIGlyGluTyrSerIleSyrLeuAspGlyLeuValIlePheSer 125
 DB 4330 AAGACGACCTCGACGACGAGG-----AAATTCGACGCGCGGATACCTCACTCA 4380
 QY 126 ValAspAlaGluIleArgGluIleTyrAsnGluGluCysProValAlaIThrAspValSer 145
 DB 4381 ACATGCGGTCAATTTCAGAGGACCGTGGCGCCAGATACCGGAGCATGATTCACAGC 4440
 QY 146 ValProLeuAspGlyArgGlnTrpSerLeuSerIlePheSerPheProMetPheArgThr 165
 DB 4441 CTACCGGTGATGGCGGAGCTGCTCTCTAGTGAATCTCCGTTCTTCAGGAT 4500
 QY 166 AlaTyrValAlaValAlaAsnValGluAsnIleSglumetSerLeuAspValAlaSer--- 184
 DB 4501 CCGTGTGTTGTCATCACCACACGACGACGACGAA-----GTGCAAGTACGAAAGCC 4554
 QY 185 AspLeu-----IleGluTrpLeuAsnAsnLeuAlaAspTrpArgTyrValVal 200
 DB 4555 GATCTGATGCGTTCGCGAAGCATTCG---AACACAGAGACGACTGACCGAAGCGAGC 4611
 QY 201 AspSerIleGluTrpIleAsnPheThrAsnAspThrIleTyrTyrValArgIleArgVal 220
 DB 4612 ---TACCAAGTTGGGCGCAAGTCGGGAC-----GTGTTTACATGATGTCGTC----- 4656
 QY 221 LeuArgProThr-----TyrAspValProAspProThrGlu---GlyLeuValArg 236
 DB 4657 ---CCGACCGAAGCGCTGACGACGCTACACCCCGACTCACTACGGGTATTCAGGG 4710
 QY 237 ThrValSerAspTyrArgGluThrTyrIleAlaIleThrCysGluAlaAsnMetProThr 256
 DB 4711 TTATCGAAGATTACCTGTGACATTCAGAGCGGCTACAGGATCTCAACGACCCACT 4770
 QY 257 LeuValAspGlnGlyPheTrpIleGlyGlyIleGlyValAlaLeuThrProThrSerLeuPro 276
 DB 4771 CTCGTGATCAGGAGAGGCGGCTGATCGCCAGTTC-----CAACCGCAAGAAACAC 4824
 QY 277 GlnTyrAspAlaSerGluAlaTyrAlaLeuIleHisThrLeuThr-----290
 DB 4825 CAGACGAGAACCCGCGCATATGATGCCGCGACACCCAAACGGCGGACACTTCACACTC 4884
 QY 291 ---PheAlaArgProSerSerAlaAlaIleAlaLeuAlaPheValTrpAlaGlyLeuProGln 309
 DB 4885 GCGGTCACAGGCGCGAATCACAATGACATGACGATGACG---ATCGGCGACAGGTGACGTTTC 4941
 QY 310 GlyGlyThrAla---ProAlaGlyThrProAlaTrpGluGlnAlaSerGlyGlyTyr 328
 DB 4942 GGGGCGCACCAATCCGCTACCCACGCTGTCATGGGCGCGATCGCGAGTCGCGGACAG 5001
 QY 329 LeuThrTrpArgHisAsnGlyThrThrPheProAlaGly-----SerValSerTyr 345
 DB 5002 CTGGTGTTCAGACTGCGCAACTGACATTCGACCTCGGAACACACATCACCATACGACAC 5061
 QY 346 ValLeuProGluGlyPheAlaLeuGluArgTyrAspProAsnAspGlySerTrpThrAsp 365
 DB 5062 ACGCTGCACACAGGCGTGGTG-----ACGGGAATGTGCAATTC 5100
 QY 366 PheAlaSerAlaGly---AspThrValThrPheArgGlnValAlaValAspGluValVal 384
 DB 5101 ACACCGACGACAGGAGGACGACACCGTGCAC-----GTGGAC-----5136
 QY 385 ValThrAsnAsnProAlaGlyGlyGlySerAlaProThrPheThrValArgValProPro 404
 DB 5137 -----GCGGAGCGCGACTGACCTGTCGAGCGAATTTGGAGCC 5178
 QY 405 SerAsnAlaTyrThrAsnThrValPheArgAsnThrLeuLeuGluThrArgProSerSer 424
 DB 5179 TCGGAAGTGAATCTGACGAGCATC-----5202


```

QY 425 ArgArgLeuGluLeuPrometProProAlaAspPheGlyLthrValAlaAsnAsnPro 444
DB 5203 -----ACTCATCAAGATTCACCAACAGACATGACCAATGATGACGACGCC 5256
QY 445 LysIleGluGlnSerLeuLeuLysGluThrLeuGlyCysArgLeuValHisSerLysMet 464
DB 5257 AAGACCATTCAGTCCACATCAACAGCAAGCAAGGCTTTATATG---CCCTGAGGGCG 5313
QY 465 ArgAsnProValPheGlnLeuThrProAlaSerSerPheGlyAlaValHisPheAsnAsn 484
DB 5314 TTCACACCCGTTTCGAATGACAAATGACATGCGACGCTTATGACCGGTCGATGAAAGACA 5373
QY 485 ProGlyTyrGluArgThrArgAspLeuProAspTyr-----ThrGlyLeuArg 500
DB 5374 CCG-----AGACACA---ACTGTGTAGATTACCAACCGGCAATTTGGTGCATCCAG 5421
QY 501 AspSerPheAspGlnAsnMetSerThrAlaValAlaHisPheArgSerLeuSerHisSer 520
DB 5422 GATACCATGACAGCAACTTCGCGATAGGCGTTGCCGCGATGACCGGTATGTCTACATCA 5481
QY 521 CysSerIleValThrLysThrArgGlnGlyTyrGluGlyValThrAsnValAsnThrPro 540
DB 5482 ACCGTACCTTACTTCAAGGTGTTCCGACGCTTCGAAGCGATTACCGCGGAGGGAGCCCT 5541
QY 541 PheGlyGlnPheAlaHisAlaGlyLeuLeuLysAsnGluLthrLeuGlyLysLeuAlaAsp 560
DB 5542 TGGGGCCCTTGGCCAGTGGACACCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5601
QY 561 AspLeuAlaThrArgLeuThrGlyValTyrProAlaThrAspAsnPheAlaAlaVal 580
DB 5602 ACTTGACCGAGTTCGACCCATTCGCGATCCCGAAGCATACAAAGCATTCGGGGCCCTA 5661
QY 581 SerAlaPheAlaAlaAsnMetLeuSerSerValLeuLysSerLthrLthrSerSerIle 600
DB 5662 TTCGCGATGTGGCCACAGACATACCCGAGATTCCTGC-----TAT 5703
QY 601 IleLysSerValGlyLthrAlaValAlaGlyAlaAlaGlnSerGlyLeuAlaLysLeuPro 620
DB 5704 GTGGCATCA-----GCAGCGGAGTGGCGAATGGCGGTGACGCGACTCATAGAG 5751
QY 621 GlyLeuLeuMetSerValProGlyLysIleAlaAlaArgValArgAlaArgAlaArg 640
DB 5752 ACCGCGACCGAGAGTGTCTCGAATTCACCTCACCTCGAGAGGCGGCAAGAGAGAGAG 5811
QY 641 Arg-----ArgAlaAlaArg 645
DB 5812 CGTGTGGCGGAGATCGCTCGAGAGCCCGC 5841
RESULT 12
ARI59969 6534 bp DNA linear PART 17-OCT-2001
LOCUS DEFINITION Sequence 4 from patent US 6251654.
ACCESSION ARI59969
VERSION ARI59969.1 GI:16222854
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 6534)
AUTHORS Gordon,K.Heinrich, and Hanzlik,T.Nelson.
TITLE Modified small RNA viruses
JOURNAL Patent: US 6251654-A 4 26-JUN-2001;
FEATURES
Location/Qualifiers
source 1..6534
BASE COUNT 1667 a 1782 c 1747 g 1338 t
ORIGIN
Alignment Scores:
Pred. No.: 1.56e-09 Length: 6534
Score: 294.00 Matches: 184
Percent Similarity: 37.50% Conservative: 74

```

```

Best Local Similarity: 26.74% Mismatches: 281
Query Match: 8.71% Indels: 151
DB: 6 Gaps: 35
US-09-677-653a-50 (1-647) x ARI59969 (1-6534)
QY 2 GlyAspAlaGlyValAlaSerGlnArgProHisAsnArgGlyLthrArgAsnValArg 21
DB 4012 GGGCAACACAGGGCGGGCAACACACACACACACACACACACACACACACACACACAC 4071
QY 22 ValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArgArgLthrGlyArg 41
DB 4072 CTTCCCGCCCTAGTCCACCGCATT----- 4095
QY 42 GlnValSerProProAspAsnPheThrAlaAlaGlnAspLeuAlaGlnSerLeuAsp 61
DB 4096 ---ACGACCCCGCCAGATGGCCGACACCCGAC------CAC 4128
QY 62 AlaAsnThrValThrPheProAlaAsnIleSerSerMetProGluPheArgAsnThrAla 81
DB 4129 GCGACACCCGGGTCAA-CCGGGTG-----GACACGCGTGGGGGCT 4172
QY 82 LysGlyLysIleAspLeuAsp-----SerAspSerIleGlyTyrPyrPheLysTyr 98
DB 4173 AAGGCAAGCGATGATGAGAGTCCCATGGCAGCGCATCAAGCGCATTCACGACTAT 4232
QY 99 LeuAspProAlaGlyLthrGlnSerAlaArgAlaValGlyLysLysLysLysLysLys 118
DB 4233 CTAGACCCCGAGCGAGAAATACAAAGACGAGCTGACGACGGG-----AAATATCC 4283
QY 119 AspGlyLeuValLysPheSerValAspAlaGlnLthrGluLthrLysAsnGluGlyCys 138
DB 4284 GACGGCGCATCTCATGCAACATCGCGGATTCGAGGAGCGCGTGGCGCCGACATAC 4343
QY 139 ProValValThrAspValSerValProLeuAspGlyArgGlnThrSerLeuSerIlePhe 158
DB 4344 CCGGACGTGAATTCAGACGCTACCGCTGAGTGGGGGAGACCTGGCTCTAGAGATG 4403
QY 159 SerPheProMetPheArgThrAlaThrAlaValAlaAlaAsnValGlnLysGluMet 178
DB 4404 CATCTCCGTTCTTCAAGCGATCCGTTGTTTCATCACCAACCGACGACGAGAA---- 4460
QY 179 SerLeuAspValAlaAsn---AspLeu-----IleGluThrLeuAsnAsnLeu 193
DB 4461 ---GTCAAGTACGACAAACCGCGATGTGATGCTGCGGACGACATTTG---AACACAGG 4514
QY 194 AlaAspTyrArgTyrValValAspSerGluGlnThrIleAsnPheThrAsnAspThrThr 213
DB 4515 ACAGACTGGACCGAAGCGAGC---TACCCAGATTGGCGCAAGTCCGGAGAC-----GTG 4565
QY 214 TyrTyrValArgIleArgValLeuArgProThr-----TyrAspValProAspPro 230
DB 4566 TTTTACATGTGCTC-----CCGACCGAAGCGCTGACGAGCATACACCCCGC 4613
QY 231 ThrGlu---GlyLeuValArgThrValSerAspTyrArgLeuThrTyrLysAlaIleThr 249
DB 4614 ACTCAACGTGGGTATCAAGGTTACTCGAGAGTACTCGTGCATCGACGAGCGGCTCAAA 4673
QY 250 CysGluAlaAsnMetProThrLeuValAspGlnGlyPheThrPheGlyGlyGlnTyrAla 269
DB 4674 GCGTACTTCAAGCAACCCACTCGTAATCGAGGAGTGGCGGTGATCCGCAATTC--- 4730
QY 270 LeuThrProThrSerLeuProGlnThrAspValSerGlnAlaThrAlaLeuHisThrLeu 289
DB 4731 ---CAACCGACAAACACACACAGAAAGAACCCGAGCATGTACTCCGACCAACCAA 4787
QY 290 Thr-----PheAlaArgProSerSerAlaAlaAlaLeuAlaPhe 302
DB 4788 ACGGCGGGAAGCTTACAGTCCGCGGTTCAGGCGCCCACTACACATTCGAGTACGAG--- 4844
QY 303 ValThrPheGlyLeuProGlnGlyLthrAla---ProAlaGlyThrProAlaThrGlu 321
DB 4845 ATCGGGGACCAAGTGTAGTTGCGGGGCGACGAAATCCCTACCCACAGCGGTGCATGGGG 4904

```

```

QY 322 GlnAlaSerSerGlyGlyTyrLeuThrProArgHisAsnGlyThrThrProAlaGly 341
DB 4905 CCGATCCCGGAGTGGGGAGCTGGTGTCCAGACTGCGAAGCTGACATTCGACGTCGA 4964
QY 342 -----SerValSerTyrValLeuProGluGlyPheAlaLeuGluArgTyrAspPro 358
DB 4965 AACACATATCCATCAAGACCGACGCTCCACCAAGGCGGTG----- 5006
QY 359 AsnAspGlySerTyrThrAspPheAlaSerAlaGly---AspThrValThrPheArgGln 377
DB 5007 ---ACGGGATGTGGCAATTCACAGCCAGCAGGAGCAGACCGTGACC----- 5054
QY 378 ValAlaValAspGluValValThrAsnAspProAlaGlyGlySerAlaProThr 397
DB 5055 -----GTGGAC-----GCGGAGCG----- 5069
QY 398 PheThrValArgValProProSerAsnAlaTyrThrAsnThrValPheArgAsnThrLeu 417
DB 5070 ---ACTGTACGCGTTCGGACGGAATTTGGACGCGCTCGGAATCTGACAGCATCAA 5126
QY 418 LeuGluThrArgProSerSerArgArgLeuGluLeuProMetProProAlaAspPheGly 437
DB 5127 CTCATATCAAGATTCACACAC-----GACATGAACCAATGATGACAGGC 5171
QY 438 GlnThrValAlaAsnAsnProGlyIleGluGlnSerLeuLeuGly---GluThrLeuGlyCys 457
DB 5172 -----AACGCCAAGACCATTCAGTCCATCAACCAAGCAGGATATATAT 5219
QY 457 sTyrLeuValHisSerLysMetArgAsnProValPheGlnLeuThrProAlaSerSerPh 477
DB 5220 CCG-----AGCGCTCAATCCGCTTTCGAATGACAAAT---GCGAGCTTTA 5263
QY 477 eGlyAlaValSerPheAsnAsnProGlyTyrGluArgTyrArgAsp---LeuProAspT 496
DB 5264 TGGACGCGTCGAT-----GAAAGACACGAGAGCAACTGTGTAGATT 5305
QY 496 yThrGly-----IleArgAspSerPheAspGlnAsnMetSer----- 508
DB 5306 ACACCGGGCAATTGTGCTACTCCAGATACATCGACACACTTCGCGATAGCTGCC 5365
QY 509 -----ThrAlaValAlaHis-PheArgSerLeuSerHisSerCysSerIleValThr 525
DB 5366 GTGATGACCGGTATGTCTATCAACCTACCTACTTGCAGGTTCGCGAGCTTCG 5425
QY 526 LysThrTyrGln-GlyTyrPgluGlyValThrAsnValAsnThrProPheGlyGlnPheAl 545
DB 5426 AACGATACCGGGAGGAGGAG-----CCTTGGGGCCCTTCGCG 5464
QY 545 aHisAlaGlyLeuLeuLysAsnGluGluIleLeuCysLeuAlaAspSerLeuAlaThrAr 565
DB 5465 TAGTGGACACCTCCGAGAGAGCGAGTGGCGTAAAGTGGCTGGAACCTGGACCGATCT 5524
QY 565 gLeuThrGlyValTyrProAlaThrAspAsnPheAlaAlaValSerAlaPheAlaAl 585
DB 5525 GCACCCATTCGCAATACCGGAGAGATCAACGATTCGGGGCCCTATTCGAGATGGCG 5584
QY 585 aAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIleIleLysSerValG 605
DB 5585 CAAGACCATAGCCAGATACCTCGC-----TATGTGCGATCA----- 5621
QY 605 yGluThrAlaValAlaGlnInsSerGlyLeuAlaLysLeuProGlyLeuLeuMetSe 625
DB 5622 -----GCGACCGGAGTGGCGAATGCGTGACGCGACTGCATAGACCGCGAGAG 5674
QY 625 rValProGlyLysIleAlaAlaArgValArgAlaArgArgAlaArgArg----- 641
DB 5675 TGTAGCTCGAATTCACCTCGAGAGCGGCAACGAAGCGGAGCTGTGTGCGGAT 5724
QY 642 -----ArgAlaAlaArg 645
DB 5735 CGCTCGAGGAGCCCGC 5750

```

```

RESULT 13
AX028953 6536 bp DNA linear PAT 16-SEP-2000
LOCUS AX028953
DEFINITION Sequence 4 from Patent WO9746666.
ACCESSION AX028953
VERSION AX028953.1 GI:10189994
KEYWORDS
SOURCE
ORGANISM Nudarellia capensis beta virus.
Nudarellia capensis beta virus
viruses; ssRNA positive-strand viruses, no DNA stage; Tetraviridae;
Betatetraviruses.
REFERENCE
1 (bases 1 to 6536)
AUTHORS Gordon, R.H. and Hanzlik, T.N.
TITLE Modified small rna viruses
JOURNAL Patent: WO 974666-A 4 11-DEC-1997;
COMM SCIENT IND RES ORG (AU) ; GORDON KARL HEINRICH (AU) ; HANZLIK
TERRY NELSON (AU)
FEATURES
source location/Qualifiers
1..6536
BASE COUNT 1667 a 1782 c 1747 g 1338 t 2 others
ORIGIN
/d_b_xref="taxon:85652"
Alignment Scores:
Pred. No.: 1.56e-09 Length: 6536
Score: 294.00 Matches: 184
Percent Similarity: 37.50% Conservative: 74
Best Local Similarity: 26.74% Mismatches: 281
Query Match: 8.71% Indels: 151
DB: Gaps: 35
US-09-677-653a-50 (1-647) x AX028953 (1-6536)
QY 2 GlnAspAlaGlyValAlaSerGlnArgProHisAsnArgGlyThrArgAsnValArg 21
DB 4014 GGGCAACCGAGGGCGCGCCCAACACACAAACAGCTGACGACGAGGCGGGGTAAAA 4073
QY 22 ValSerAlaAsnThrValThrValAsnGlyArgArgAsnGlnArgArgGlyThrGlyArg 41
DB 4074 CTTCCCGCCGTAAGTCGACCGCATP----- 4097
QY 42 GlnValSerProProAspAsnPheThrAlaAlaAlaGlnAspLeuAlaGlnSerLeuAsp 61
DB 4098 ---ACGACCGCGCGAGATGGCCGACCCGAC-----CAC 4130
QY 62 AlaAsnThrValThrPheProAlaAsnIleSerSerMetProGluPheArgAsnTrpAla 81
DB 4131 GCGAAGACCGGGGTCAA-CCGGGTGCG-----GACACGCGTGAAGGGGCT 4174
QY 82 LysGlyLysIleAspLeuAsp-----SerAspSerIleGlyTyrPheLysTyr 98
DB 4175 AAGGCAAGCGATGTGAGAGTCCATGCGACGACCATCAAGCGCATTCACGACTAT 4234
QY 99 LeuAspProAlaGlyValThrGluSerAlaArgAlaValAlaGlyLysTyrSerIlePro 118
DB 4235 CTAGACCCGAGCGAGATCAACAGACGACCTGACACAGCG-----AAATTCCTC 4285
QY 119 AspGlyLeuValLysPheSerValAspAlaGluIleArgGluIleTyrAsnGluGlyCys 138
DB 4286 GACGGCGCATCTCAGTCAACATCGGATTCGAGGAGACCGTGGCGCCAGATAC 4345
QY 139 ProValValThrAspValSerValProLeuAspGlyArgGlnTyrPheSerIlePhe 158
DB 4346 CCGGGACTGAATTCCTACAGCGCTACCGCTGATGGCGGAGACCTGCTACTAGTATG 4405
QY 159 SerPheProMetPheArgThrAlaTyrValAlaValAlaAsnValGluAsnLysGluMet 178
DB 4406 CATCTCCCGTTCTTCAGGCAATCCGTTGTTCATCAACACACACACACACAGGAA--- 4462
QY 179 SerLeuAspValValAsn---AspLeu-----IleGluTyrPheAsnAsnLeu 193
DB 4463 ---GTCAAGAGCAAGACGCGCATCTGATGCTTCCGGAACGATGG---AACAAACAGG 4516

```

```

QY 194 AlaAspTrpArgTyrValValAspSerGluInTrpIleAsnProPheGlnPheAla 213
    |||||
DB 4517 ACGGACTGACCGAAGCCAGC---TACCCAACTTGGCGCAATGGGAAC-----GTG 4567
QY 214 TyrTrpValArgIleArgValIleuArgProThr-----TyrAspValProAspPro 230
    |||||
DB 4568 TTTTACATGATGTCGTC-----CCGACCGAAGCGGTGACGAGACGTACACACCCCG 4615
QY 231 ThrGlu---GlyLeuValArgThrValAspSerPylrArgLeuThrTyrLeuAlaIleThr 249
    |||||
DB 4616 ACTGACAGTGGTGTATCAGGGGTATCGAGATTACCGTCTGATCTGACATCGACGGCGTCA 4675
QY 250 CysGluAlaAsnMetProThrLeuValAspGlnGlyPheThrIleGlyGlnGlyTrpAla 269
    |||||
DB 4676 GCGTACTTCACAGCACCACCTCTGCTGTATCAGGAGTGGCGGTGATCGCCAGTTC--- 4732
QY 270 LeuThrProThrSerLeuProGlnTyrAspValSerGluAlaIleThrLeuAlaIleThr 289
    |||||
DB 4733 ---CAACCGGACAAAGAACACAGAGAGAACCCGACATGATACCGGGACACCCAA 4789
QY 290 Thr-----PheAlaArgProSerSerAlaAlaIleuAlaIleuAlaPhe 302
    |||||
DB 4790 ACGGGCGGAACGTTCACGTCTGCGGTTCAGGGCCGAACTACACATTGACGATGACG--- 4846
QY 303 ValTrpAlaGlyLeuProGlnGlyTrpAla---ProAlaGlyThrProAlaIleuAla 321
    |||||
DB 4847 ATCGGGACCAAGTCGATTCGGGGCGCAGCAATCCCGCTACCCAGGTGTCTGATGGGG 4906
QY 322 GlnAlaSerSerGlyTyrLeuThrTrpArgHisAsnGlyThrPheProAlaGly 341
    |||||
DB 4907 CCGATCGCGGAGTCCGGGAGCTGTCTCCAGCTCGCAACCTGACATTCGACGTGGGA 4966
QY 342 ---SerValSerTyrValLeuProGlnGlyPheAlaIleuAlaIleuAlaIleuAla 358
    |||||
DB 4967 AACACATTCACATCACAGCACGCTCCACACAGGCTCGGTG----- 5008
QY 359 AsnAspGlySerTrpTrpAspPheAlaSerAlaGly---AspThrValIlePheArgGln 377
    |||||
DB 5009 ---ACGGGAATGTGGCAATTCACAGCCAGACAGCGGAGGACACCGGAGC----- 5056
QY 378 ValAlaValAspGluValValIleThrAsnAsnProAlaGlyGlySerAlaIleuAla 397
    |||||
DB 5057 ---GTGAC-----CGCGGAGCG----- 5071
QY 398 PheThrValArgValProProSerAsnAlaTyrThrAsnThrValPheArgAsnThrLeu 417
    |||||
DB 5072 ---ACTGTACGCGCTCGAGCGAATTGGACGCTCGAAGCTGATCTCGAGACATCAA 5128
QY 418 LeuGluThrArgProSerArgArgLeuGluLeuProMetProProAlaAspPheGly 437
    |||||
DB 5129 CTCATTCAGATTCACCAAC-----GACATGAACCCAAATGATGACAGCG 5173
QY 438 GlnThrValAlaAsnAsnProLysIleGluInSerLeuLys---GluThrLeuGlyC 457
    |||||
DB 5174 ---AACGCCAAGACCATTCAGTCCCACTCAACGAAGAGAGCGCTTTATATG 5221
QY 457 sTrpLeuValHisSerLysMetArgAsnProValPheGlnLeuThrProAlaSerSerPhe 477
    |||||
DB 5222 CCTG-----AGCGCTCAATCGGTTCGAAATGACAAAT---GCGACGCTTAA 5265
QY 477 eGlyAlaValSerPheAsnAsnProGlyTyrGluArgThrArgAsp---LeuProAspTr 496
    |||||
DB 5266 TGGACCGGTCGAT-----GAGACACCGAGGACACTGTGTGATG 5307
QY 496 TyrThrGly-----IleArgAspSerPheAspGlnAsnMetSer----- 508
    |||||
DB 5308 ACACGGGGAATGTGGCACTCGAAGATACCATGACACGCAACTTCGCGGATAGGTTGCC 5367
QY 509 ---ThrAlaValAlaHis-PheArgSerLeuSerHisSerCysSerIleValThr 525
    |||||
DB 5368 GTGATGACCGGTATGTCTACATCAACGCTACCTGACAAAGGTGTCCGACCGTTGG 5427

```

```

QY 526 LysThrTyrGln-GlyTyrPheGluGlyValThrAsnValAsnThrProPheGlnPheAla 545
    |||||
DB 5428 AAGCGATACCGCGCGGAGGAGGAG-----CCTTGGGGCCCTTGGC 5466
QY 545 HisAlaGlyLeuLeuLysAsnGluGluIleuLysLeuAlaAspAlaIleThrArg 565
    |||||
DB 5467 TAGTCCGACACCTCCGAAGAGACGAGCTGGCTTAACATGCTCGCACTGGACCGATCT 5526
QY 565 GLeuThrGlyValTyrProAlaThrAspAsnPheAlaAlaValSerAlaPheAlaI 585
    |||||
DB 5527 GCACCAATTCGATACCGCGGAACAGATACACAGGATTCGGGGCCCTATTCGCGATGTGGC 5586
QY 585 AsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIleIleLysSerValG 605
    |||||
DB 5587 CAAGACCATACCGACGATACCTCC-----TATGTGCATCA----- 5623
QY 605 YGluThrAlaValaGlyAlaAlaGlnSerGlyLeuAlaLysLeuProGlyLeuLeuMetSe 625
    |||||
DB 5624 ---GACCCGAGTGGCGAATTCGCTGACGAGTCTGATAGAGACCGGACCGAGAG 5676
QY 625 rValProGlyLysIleAlaAlaArgValArgAlaArgAlaArg----- 641
    |||||
DB 5677 TGTAGCCCTCGAATTCACACTCGAGAGAGCGGCAACGAAGACGAGACTGTGGCGGAT 5736
QY 642 ---ArgAlaAlaArg 645
    |||||
DB 5737 CGCTCGAGAGACCCCGC 5752
RESULT 14
LOCUS AX028972/c 6536 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 23 from Patent WO9746666.
ACCESSION AX028972
VERSION AX028972.1 GI:10190010
KEYWORDS
SOURCE Nudaurrella capensis beta virus.
ORGANISM Nudaurrella capensis beta virus
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tetraviridae;
Betatevirivirus.
REFERENCE 1 (bases 1 to 6536)
AUTHORS Gordon,K.H. and Hanzlik,T.N.
TITLE Modified small rna viruses
JOURNAL Patent: WO 9746666-A 23 11-DEC-1997;
COMMUN SCIENT IND RES ORG (AU) ; GORDON KARL HEINRICH (AU) ; HANZLIK
TERRY NELSON (AU)
FEATURES
source 1..6536
location/Qualifiers
BASE COUNT 1338 a 1747 c 1782 g 1667 t 2 others
ORIGIN
Alignment Scores:
pred. No.: 1.56e-09 Length: 6536
Score: 294.00 Matches: 184
Percent Similarity: 37.50% Conservative: 74
Best Local Similarity: 26.74% Mismatches: 281
Query Match: 8.71% Indels: 151
Gaps: 35
US-09-677-653A-50 (1-647) x AX028972 (1-6536)
QY 2 GlyAspAlaGlyValAlaSerGlnArgProHisAsnArgArgGlyThrArgAsnValArg 21
    |||||
DB 2523 GGGCAACAGGCGCGCGCAACACAAACAAAGCTCGACGAGGAGGCGGTCTAAAA 2464
QY 22 ValSerAlaAsnThrValThrValaLysGlnArgArgAsnGlnArgArgThrGlyArg 41
    |||||
DB 2463 CTTCCTCCCGTGTGTCGACCGAT----- 2440
QY 42 GlnValSerProProAspAsnPheThrAlaAlaGlnAspLeuAlaGlnSerLeuAsp 61
    |||||
DB 2439 ---ACCAAGCCCGCAGATGGCCGAACCCCGAAC-----CAC 2407

```

QY 62 AlaAsnThrValIThrPheProAlaAsnIleSerSerMetProGluPheArgAsnTrpAla 81
 DB 2406 GCGAACAACCCGGGTCAA-CCGGCGCTC-----GACACGGGTGAGGGGCT 2363
 QY 82 LysGlyLysIleAspLeuAsp-----SerAspSerIleGlyTrpTrpPheLysTy 98
 DB 2362 AAGGCAAGAGATGATGAGAGATGCCATGGCAGGACATCAAGGATGATTCACGACTAT 2303
 QY 99 LeuAspProAlaGlyAlaThrGluSerAlaArgAlaValGlyGluTrpSerLysIlePro 118
 DB 2302 CTAGACCCCGACGAGAAATACAGACAGCCTGAGACGAGCGG-----AAATATCCC 2252
 QY 119 AspGlyLeuValLysPheSerValAspAlaGluIleArgGluIleTrpAsnGluGly 138
 DB 2251 GACGGGCGCATCTCACTGATCAACATCGGCTCAATTTGAGGAGCGCTGGCGCCAGATAC 2192
 QY 139 ProValValThrAspValSerValProLeuAspGlyArgGluTrpSerLeuSerIlePhe 158
 DB 2191 CCGGAGCTGAATCTACGACGCTACCGCTGATGGGAGCGGCTGCTACTAGTGATG 2132
 QY 159 SerPheProMetPheArgThrAlaIleValAlaValAlaAsnValGluAsnLysGluMet 178
 DB 2131 CATCTCCCTCTCTTCAGGATCCGTTGTTCTCATCACACACACGACGACGAGAA--- 2075
 QY 179 SerLeuAspValValAsn---AspLeu-----IleGluTrpLeuAsnAsnLeu 193
 DB 2074 ---GTGAAGTGCAGAACCCGATCTGATGCTTGCAGACGATGG---AACAAACAGG 2021
 QY 194 AlaAspTrpArgTrpValValAspSerGluInTrpIleAsnPheThrAsnAspThrThr 213
 DB 2020 ACGGACTGGACGACGACGACG---TACCCAGTTGGCGGCAATCGGGAAC-----GTG 1970
 QY 214 TyrTrpValArgIleArgValLeuArgProThr-----TyrAspValProAspPro 230
 DB 1969 TTTTACATGTCGCTC-----CCGACCGAAGCGCTGACGACGACGACGACCCCG 1922
 QY 231 ThrGlu---GlyLeuValArgThrValSerAspTrpArgLeuThrTrpLysAlaIleThr 249
 DB 1921 ACCTCACTGGGTGTATCAGAGGTACTCGAGAGTTCACCGTGTGACATCGAGCGGCTCACA 1862
 QY 250 CysGluAlaAsnMetProThrLeuValAspGluGlyPheTrpIleGlyGluTrpAla 269
 DB 1861 GCCTACTTCAACGACCCACTCTCGTAATCAGGAGGTGGGTGATCGCGCAGTTTC--- 1805
 QY 270 LeuThrProThrSerLeuProGluInTrpAspValSerGluAlaIleThrLeu 289
 DB 1804 ---CAACCGACAAAGAACACACGAGAGAACCCGACATAGTACCCGACCCCA 1748
 QY 290 Thr-----PheAlaArgProSerSerAlaAlaAlaLeuAlaPhe 302
 DB 1747 ACGGGGGAACGTTACAGCTCGCGGCTACGAGCCGAACTACACATTTGACGATGAG--- 1691
 QY 303 ValTrpAlaGlyLeuProGluGlyTrpAla---ProAlaGlyTrpProAlaTrpGlu 321
 DB 1690 ATCGGGGACACGATCGAGTTCGGGGCGCAGCAATCCCGCTACCGAGGTGATGGGG 1631
 QY 322 GlnAlaSerSerGlyLysTrpLeuThrTrpArgHisAsnGlyThrThrPheProAlaGly 341
 DB 1630 CCGATCCCGAGTGGGGGAGCTGTGTTCACAGACTCGCAACCTGACATTCGACGTGCGA 1571
 QY 342 -----SerValSerTrpValLeuProGluGlyPheAlaLeuGluIleArgTrpAspPro 358
 DB 1570 AACACATACACATCAACGACGACGACGCTGCCACCGGCTCGT----- 1529
 QY 359 AsnAspGlySerTrpThrAspPheAlaSerAlaGly---AspThrValThrPheArgGln 377
 DB 1528 ---ACGGAATGTCGCAATTCACAGCCAGACGAGGACGACCGGAGC----- 1481
 QY 378 ValAlaValAspGluValValValThrAsnAspProAlaGlyGlyLysSerAlaProThr 397
 DB 1480 ---GTGAC-----GCGGACGCG----- 1466

QY 398 PheThrValArgValProProSerAsnAlaTyrThrAsnThrValPheArgAsnThrLeu 417
 DB 1465 ---ACTGTACGCTTTCGAGCCGATTTGGACGCTCGGACACTGAATCTGACGACTCAA 1409
 QY 418 LeuGluThrArgProSerSerArgArgLeuGluLeuProMetProProAlaAspPheGly 437
 DB 1408 CTCATCAAGATTCACCAAC-----GACATGAACCAATGATGACGAGC 1364
 QY 438 GlnThrValAlaAsnAsnProLysIleGluInSerLeuLeuLys-GluThrLeuGly 457
 DB 1363 -----AACCCCAAGACCATTCAGTTCCAACTACGAAAGAGGCGATTATATG 1316
 QY 457 sTrpValHisSerLysMetArgAsnProValPheGluInLeuThrProAlaSerSerPhe 477
 DB 1315 CCTG-----AGCGTCAATCCGTTGTGAAATGACAT- GCGAGCTTTA 1272
 QY 477 eGlyAlaValSerPheAsnAsnProGlyLysArgTrpArgAsp---LeuProAsp 496
 DB 1271 TGGACCGGTGAT-----GAAAGACCGAGACGACGACGATTTAT 1230
 QY 496 yThrGly-----IleArgAspSerPheAspGlnAsnMetSer----- 508
 DB 1229 ACACCGGCAATGTGTGACCTCCAGAGATACATCGACAGCAACTTCGAGATAGTTGCC 1170
 QY 509 -----ThrAlaValAlaHis-PheArgSerLeuSerHisSerCysSerIleValThr 525
 DB 1169 GTGATGACCGGATGTCTATCATCAACGTTACCTTTCGAAAGTTCGAGCTTGG 1110
 QY 526 LysThrTyroGln-GlyTrpGluGlyValThrAsnValAsnThrProPheGlyGlnPheAl 545
 DB 1109 AAGCATACCGGCGGAGGAG------CCTTGGGCGCCCTTCCG 1071
 QY 545 aHisAlaGlyLeuLeuLysAsnGluGluIleLeuGlySleuAlaAspAspLeuAlaThr 565
 DB 1070 TAGTGACACCTCCGAGAGGACGCTGGCGCTAACAGTGGCTGCACTTGGACGATCT 1011
 QY 565 gLeuThrGlyValTyrProAlaThrAspAsnPheAlaAlaValSerAlaPheAla 585
 DB 1010 GCACCATTCGCTATCCCGGAAGCATACAGGATTCGGGCGCTATTCGATGGTGGC 951
 QY 585 aAsnMetLeuSerSerValLeuLysSerGluAlaThrSerSerIleLysSerValG 605
 DB 950 CAAGACATGACCCCATATACCTCGC-----TANGTGCATCA----- 914
 QY 605 yGluThrAlaValGlyAlaAlaGlnSerGlyLeuAlaLysLeuProGlyLeuLeuMet 625
 DB 913 -----GACCGGAGTGGCGATCGGTGACGAGCTGCATGAGAGCGGACCGAGAG 861
 QY 625 rValProGlyLysIleAlaAlaValArgValArgAlaArgValArgValArgVal 641
 DB 860 TGTAGCTCGAATTCACCTCGGAGAGCGGCAACGAGAGCGAGCATGTTGGCGAAT 801
 QY 642 ---ArgAlaAlaArg 645
 DB 800 CGCTCGAGAGCGCGC 785
 RESULT 15
 AE005131
 LOCUS
 DEFINITION
 AE005131 AE004437
 ACCESSION
 VERSION
 AE005131.1 GI:10581942
 KEYWORDS
 SOURCE
 ORGANISM
 Halobacterium sp. NRC-1.
 Halobacterium sp. NRC-1
 Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 Halobacteriaceae; Halobacterium.
 1 (bases 1 to 11659)
 REFERENCE
 AUTHORS
 Ng,W.V., Kennedy,S.P., Mahairas,G.G., Bergquist,B., Pan,M.,
 Shukla,H.D., Lasky,S.R., Baliga,N., Thorsson,V., Sirogana,J.,
 Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Welti,R., Goo,Y.A.,
 Leithausen,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
 Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,

TITLE	JOURNAL	DOI	PMID	PMCID	REFERENCE	AUTHORS
Isenberger, T. A., Peck, R. F., Pohlschrod, M., Spudich, J. L., Jung, K. -H., Alam, M., Freitas, T., Hou, S., Daniels, C. J., Dennis, P. P., Omer, A. D., Ehardt, H., Lowe, T. M., Liang, P., Riley, M., Hood, L. and DasSarma, S.	Proc. Natl. Acad. Sci. U.S.A.	97 (22), 12176-12181	(2000)			
DasSarma, S.						
Genome sequence of Halobacterium species NRC-1						
20504483						
11016950						
2 (bases 1 to 11659)						
Ng, W. V., Kennedy, S. P., Mahaitas, G. G., Bergquist, B., Pan, M., Shukla, H. D., Lasky, S. R., Balliga, N., Thorsen, V., Sbrogna, J., Swartzell, S., Weir, D., Hall, J., Dahl, T. A., Welter, R., Goo, Y. A., Leitner, B., Keller, K., Cruz, R., Denson, M. J., Hough, D. W., Meddock, D. G., Jablonksi, P. E., Krebs, M. P., Angevine, C. M., Dale, H., Isenberger, T. A., Peck, R. F., Pohlschrod, M., Spudich, J. L., Jung, K. -H., Alam, M., Freitas, T., Hou, S., Daniels, C. J., Dennis, P. P., Omer, A. D., Ehardt, H., Lowe, T. M., Liang, P., Riley, M., Hood, L. and DasSarma, S.						
Direct Submission						
Submitted (14-JUL-2000)						
Roosevelt Way NE, Seattle, WA 98105, USA						
Location/Qualifiers						
1. .11659						
/organism="Halobacterium sp. NRC-1"						
/strain="NRC-1"						
/db_xref="taxon:64091"						
157. 1074						
/gene="tygc"						
/note="VNG2535G"						
157. 1074						
/gene="tygc"						
/note="tygc"						
/codon_start=1						
/transl_table=1						
/product="6-phosphogluconate dehydrogenase"						
/protein_id="AAG20606.1"						
/db_xref="GI:10581943"						
/translation="MELGVVGLGMRIVVDRCLDAGHDVVAFDIDATAREADAGA;TPADTDLALAEGLAEIRIMVLPAGRPVDADELPHILDSDDVVDGSHREDST;RAEHEHFAIDCGTSGGPAELGFGFPAWADLVPRVDVATVATGPDGDRMG;PACAGHYKVMHNGVEALVIMOTYGECELLHEGYRDDLEAVATNNNGAVIRNMILE;LCERAPREESDICTVADHVAGSGTGWTYQALEQEPVPLVSQALSERGSRSTGG;DTGFAFRILANRRKIRGFRHVAFTGDA"						
complement (1071. .1652)						
/gene="VNG2554H"						
complement (1071. .1652)						
/gene="VNG2554H"						
/note="hypothetical protein"						
/codon_start=1						
/transl_table=1						
/product="Vng2554h"						
/protein_id="AAG20607.1"						
/db_xref="GI:10581944"						
/translation="MDVAVLLFEGFDELAIIGPEYERPTAAELGGDLVSMRTLPDE;GVAARGITVAPDDVLCGVLDIVVPGGGWTDDBRPGVSEVENETPEPLDITLHGGS;ATIASVGTGAMLLADAGLLDGRPAATHASADDIAETDIDVYRDERIVDDGTIVLTAGV;TAGIDIALHYVERECSAGLAEQVATEIVASTY"						
1746. .2324						
/gene="VNG2555C"						
1746. .2324						
/gene="VNG2555C"						
/note="conserved hypothetical protein"						
/codon_start=1						
/transl_table=1						
/product="Vng2555c"						
/protein_id="AAG20608.1"						
/db_xref="GI:10581945"						
/translation="MTTIGVTLGIGVTLFLVYLITAKSSKPAEDIAQDVLEHRS;TVETDPPEPYNRSISGGGAGAVAGAAESELDEGEDGDFPDATADDDVETETE;YVEGEETIVASNEISILDAGVDEGMDIPVACQCGVCVSGRVADGEDAAEYIRHSNN;ESIDGAEIEGVMILCTAHTPAFTASFIETDTP"						
2413. .2485						
/gene="trn44"						

trna	2413. .2485	/gene="trn44"	/product="trRNA-Arg"	complement (2445. .2955)	/gene="VNG2556H"	complement (2545. .2955)	/gene="VNG2556H"	/note="hypothetical protein"	/codon_start=1	/transl_table=11	/product="Vng2556h"	/protein_id="AAG20609.1"	/db_xref="GI:10581946"	/translation="MTYDRGLLIIVAGCVLNLAVMTLHSPPIITPAADDSGGLCVL GIALVLVPEMLLGLPTMHVADHGAVCCLLTGLVTGACLMNGITAPPSESLTLYVY EAPFELVLYVVGIVIEOCLRTGHAVDSNRSQE"	complement (3099. .4019)	/gene="fepC"	/note="VNG2558g"	complement (3099. .4019)	/gene="fepC"	/note="fepC"	/codon_start=1	/transl_table=11	/product="ferric enterobactin transport protein"	/protein_id="AAG20610.1"	/db_xref="GI:10581947"	/translation="MTROPANDATGTDPDRGLSGLDVLVRI PGGDPEV IDGESVTP VDVVALIIPNGSGSKSTLLNALANKHSPADGTVLLDHT IDELSKEIARLLGISOE NYPGGSISRELVERGRYPHOGEFDLSLDADRAAVTDALRVAGVDHLSGRDVGSLSGG OKYOLVIAALAQDIDVLLIDEPTEFLDPHOLEVMDI IQALRDPDHDTVVVVLHDI QAARYDHTIALKDGITIHARGOPADVVRKETLRVFIEDVAVSTTSHQMVGVPTEALH DEGTDHENAIPGSRPTPTETADGCTPPAE"	complement (4023. .5102)	/gene="yJmd2"	/note="VNG2560G"	complement (4023. .5102)	/gene="yJmd2"	/note="yJmd2"	/codon_start=1	/transl_table=11	/product="ferrichrome ABC transporter permease protein"	/protein_id="AAG20611.1"	/db_xref="GI:10581948"	/translation="MTANTDASATRLATLAAGSLVLYATLLQVSGFAPPIADPA MHWDEPELSSAGVRAAFILGGOLPEMLSRQVYVNNIRLPRLVGLIVGANLAVSG ALFOIYTRRELASPYLIGVNSAGMNVLLTLTISGLIPWLLAALGAVAFILVYA TAMKNGTNRYRVLVAGVYVSYFGSVQALFFPIDNLGVNSAOTWLSGSLGTDMAQ VRIALPEVLTSLALALVTNELDVLLDEETRESLSGMEVEKVRFGVAGTALSTAAI AAVAGLIEFGLLIPHVRNIVGNNTRDDLLACCVLGPALLVGADVGARLADPOLPVPI GIITGVGSGPYFLYLMRRTENLGGI"	complement (5109. .6263)	/gene="VNG2562H"	complement (5109. .6263)	/gene="VNG2562H"	/note="hypothetical protein"	/codon_start=1	/transl_table=11	/product="Vng2562h"	/protein_id="AAG20612.1"	/db_xref="GI:10581949"	/translation="MARRQLLAGASLIASLAGCISSSGPTDDGSTTOTDAGSDY IKVATIEPHTFESVPEYVAIIPVYMDIGMALIOPSATISGRAPLKFYDIALPDVA FKRAIEOLASDESQYKENEYPAADCDVHLIDPRMLKFYANWTDIDIAIESAGFP KSSSIRFASVSSVAGDPFYDLYEAFKAAETFORARFAEASIKADIMPTISGGLP ESDRPVAIIMGVDPGOFRIAPLDPDONNTLTYYRLGMDAFAGAVPDGPFGYEEL LIDVDPYIGAAVGLISQTHAIEPVDTVVEPEANNGOQLTAVENGMLVRTGQFMKP IYDLESTALAKMYPDAFGSMPSGSPETIPASENLPDRQVADINVGVL"	complement (6353. .7216)	/gene="VNG2563H"	complement (6353. .7216)	/gene="VNG2563H"	/note="hypothetical protein"	/gene="VNG2563H"	complement (6353. .7216)	/gene="VNG2563H"	/note="hypothetical protein"	/gene="VNG2563H"	complement (6353. .7216)	/gene="VNG2563H"	/note="hypothetical protein"
------	-------------	---------------	----------------------	--------------------------	------------------	--------------------------	------------------	------------------------------	----------------	------------------	---------------------	--------------------------	------------------------	--	--------------------------	--------------	------------------	--------------------------	--------------	--------------	----------------	------------------	--	--------------------------	------------------------	--	--------------------------	---------------	------------------	--------------------------	---------------	---------------	----------------	------------------	---	--------------------------	------------------------	---	--------------------------	------------------	--------------------------	------------------	------------------------------	----------------	------------------	---------------------	--------------------------	------------------------	--	--------------------------	------------------	--------------------------	------------------	------------------------------	------------------	--------------------------	------------------	------------------------------	------------------	--------------------------	------------------	------------------------------

OY	368	rAaGlyAaPThrVal---	ThrPheArgGlnIleAlaValaAspGluValValaThrAs	387
Db	1482	GTGGGGCCCAACCGTGAAGCCCGGTGGCCGGCGT	-----	1515
OY	387	nAsnProAlaGlyGlySerAlaProThrPheThrValArgValProPheSerAsnAl		407
Db	1516	-ACGCTTGCTGGGGACGAGCGGTGGCATCGACACGTGAGGTGGCGCCACACTCCGC		1574
OY	407	eTyThrAsnThValPheArgAsnThrLeuLeu	-----GluThrAr	421
Db	1575	G-----GCCGTGGGGAACACTGTGACGGGCGCATGGCGTCCAGTTGCTGGA		1622
OY	421	gProSerSerArgArg-LeuGlnLeuProMetProProAlaAspPheGlyGlnThrValA		441
Db	1623	CCCTTGAAACGACGAGACTGCAACGTCCATAACCGGCTGTGCAC	-----	1666
OY	441	lAsnAsnProLysIleGluGlnSerLeuLeuLysGluThrLeuGlyCysTyrrLeuValH		461
Db	1667	--CCACACCTCA-----		1678
OY	461	lSerLysMetArgAsnProValPheGlnLeuThrProAlaSerSerPheGlyAlaValS		481

Db	1678	-----		1678
OY	481	ePheAsnAsnProGlyTyrrGlyAlaArgThrArgAspLeuProAspTyrrThnGlyLeuArg		501
Db	1679	-----AATTCGGCCGGA-----		1691
OY	501	sPSePheAspGlnAsnMetSerThrAlaValAlaHisPheArgSerLeuSerHisSerC		521
Db	1692	AAACGTGGTGGCGCATTCACAAATGTTAAACTGGCGTTCCAATACCGCATCGATGTGT		1751
OY	521	ySerLleValThrLysThrTyrrGlnGlyTrrPheGluValThrAsnValAsnThrPro		541
Db	1752	ACCACACTGGGTCTCACGCTGGCGCTGGG-----GGTGACACTGTGTTCGTCTCTCT		1804
OY	541	heGlyInPheAlaHisAlaGlyLeuLeu-Lys-----AsnGluGlnIleLeuCysLeu		558
Db	1805	-----GCATCTCCCGGAAGGCTCTATCCAAACCCCGCGGAGGACATC-----		1847
OY	559	AlAspAspLeuAlaThrArgLeuThrGlyValTyrrProAlaThrAsp-AsnPhe-----		576
Db	1848	GCACAGACGCTCTCGAACACCGCTGGGTCCACAGTGCAGGACGAGCACTTCCCTGAGCG		1907
OY	577	-----		587
		-----AlaAlaAlaValSerAlaPheAlaAlaAsnMe		
Db	1908	TACAAACCGTCCATCGGGGCGGCGGGCGGCGGCGGTCTGCTGCGGCGGTGCGAGATCG		1967
OY	587	tLeuSerSerValLeuLysSerGlnAlaThrSerSerIleIleLysSerValGlyGluTh		607
Db	1968	GAACCTCGAAGACGGCGGAGAAAGACGACGGCTTGATTCGGAGCCGATCGCCGACAC		2027
OY	607	rAlaValGlyAlaAlaGlnSerCylLeuAlaLysLeuProGlyLeuLeuMetSerValPr		627
Db	2028	GACGTGCACTACTACGACAGTCCGAGTACGTCAAG-----		2061
OY	627	oGlyLysIleAlaAlaArgValArgAlaArgArgAlaArgArgAlaAlaArgAlaAs		647
Db	2062	-----AAGCGAGACCATCGAGGTGCGGAGCAACGAGAGATC		2099
OY	647	n 647		
Db	2100	: 2100		